

1998

Sample-Based Forest Landscape Diversity Indices.

Bernard Ross Parresol

Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_disstheses

Recommended Citation

Parresol, Bernard Ross, "Sample-Based Forest Landscape Diversity Indices." (1998). *LSU Historical Dissertations and Theses*. 6696.
https://digitalcommons.lsu.edu/gradschool_disstheses/6696

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

SAMPLE-BASED FOREST LANDSCAPE DIVERSITY INDICES

A Dissertation

**Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy**

in

The School of Forestry, Wildlife, and Fisheries

by

**Bernard Ross Parresol
B.S., Michigan State University, 1977
M.S., Louisiana State University, 1983
M.Ap.Stat., Louisiana State University, 1983
May 1998**

UMI Number: 9836897

**Copyright 1998 by
Parresol, Bernard Ross**

All rights reserved.

**UMI Microform 9836897
Copyright 1998, by UMI Company. All rights reserved.**

**This microform edition is protected against unauthorized
copying under Title 17, United States Code.**

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

©Copyright 1998
Bernard Ross Parresol
All rights reserved

DEDICATION

To my parents,
Thomas and Rita Parresol,
my wife, Lisa, and step-daughters, Sarah and Christine,
Norma Dreyfus and her Father, Norman Dreyfus,
The Many Nations Singers and all my relations,
and to my friend Charlie Thomas,
for their belief in me.

This dissertation is a manifestation of their love, frustrations, and inspirations.

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to my major professor, Dr. Quang Van Cao, for his guidance and advice throughout the study and in the preparation of this dissertation. I would like to thank the other members of my committee: Dr Jim Chambers, Dr. Brian Marx, Dr. Richard Pace and Dr. Dan Earle. I also wish to thank Dr. Vernon Wright who earlier served on my committee.

I am grateful to Joseph McCollum who assisted me in learning and using ARC/INFO and MATLAB. Appreciation is extended to Dr. Tom Dell, Dr. Charles Thomas, Dr. Margaret Devall, Dr. Greg Reams, and Dr. Paul Van Deusen for advise and encouragement during this research. I would also like to thank Dr. Harbin Li for use of his SHAPC program for generating simulated landscapes

Finally, to all my friends and relations, especially my wife Lisa, her parents Jim and Linda Morcom, and my parents Thomas and Rita Parresol, I express my sincere thanks for their moral and financial support, encouragement, and understanding during the period of my studies.

TABLE OF CONTENTS

DEDICATION	iii
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	viii
ABSTRACT	x
INTRODUCTION	1
The Landscape Scale	1
Assessing and Comparing Landscape Diversity	1
Application	5
Objectives	5
REVIEW OF LITERATURE	7
Concepts	7
Landscape Indices	8
Variance Computation from Distribution Theory	18
Variance Approximation from Nonparametric Theory	19
DERIVING NEW CONTAGION INDICES	22
Contagion Generalized	22
Geometric Distribution	23
New Contagion Index Γ_1	24
New Contagion Index Γ_2	24
SAMPLING PROPERTIES OF $\hat{\Gamma}_1$, $\hat{\Gamma}_2$, AND $\hat{\Gamma}_{2(scaled)}$	25
The O , o Notation	25
Bias	26
Variance	27
Consistency	31
Asymptotic Distribution	31
MATERIALS AND METHODS	34
Behavior of Γ_1 , Γ_2 , and RC_2	34
Increasing Gradient of Evenness	34
Same Degree of Evenness	36
Monte Carlo Investigation of Distribution	36

Data	38
Physiographic Provinces and Forest Cover Type Data	38
Study Areas	40
Hypotheses and Hypothesis Testing	43
RESULTS AND DISCUSSION	50
Behavior of Γ_1 , Γ_2 , and RC_2	50
Increasing Gradient of Evenness	50
Same Degree of Evenness	56
The Case Against Relative Contagion	63
The Meaning of Relative Contagion	67
Monte Carlo Investigation of Distribution	69
Analysis of Contagion on the Physiographic Provinces	75
Composition	78
Configuration	78
Remarks	81
Choice of Contagion Index	82
SUMMARY AND CONCLUSIONS	85
REFERENCES	89
APPENDIX A: DERIVATIONS FOR $E[X p_{ij}]$, $E[1/(X+1) p_{ij}]$, $E[X/(X+1) p_{ij}]$, AND BOUNDS FOR EQUATION (23)	94
APPENDIX B: FORTRAN CODE, EXAMPLE INPUT FILE, AND EXAMPLE OUTPUT FOR PROGRAM CONTAGION	99
APPENDIX C: FORTRAN CODE, EXAMPLE INPUT FILE, AND EXAMPLE OUTPUT FOR PROGRAM VAREST	107
APPENDIX D: FORTRAN CODE, EXAMPLE INPUT FILE, AND EXAMPLE OUTPUT FOR PROGRAM REPHIST	115
APPENDIX E: TABLES OF INDEX VALUES FROM SIMULATED LANDSCAPES	123
VITA	126

LIST OF TABLES

1. USDA Forest Service general forest cover types occurring on three physiographic provinces in the state of Alabama	41
2. Proportion of each forest cover type on each province by survey year	47
3. Illustrative example showing illogical behavior of relative contagion indices	66
4. Comparison of theoretical variance against nonparametric estimate of variance on three simulated 24 x 24 lattice landscapes of five patch types each	71
5. Contagion values, variances, and F -tests using $\hat{\Gamma}_1$ on three physiographic provinces in Alabama during three survey periods	76
6. Contagion values, variances, and F -tests using $\hat{\Gamma}_2$ on three physiographic provinces in Alabama during three survey periods	77
E1. Index values from 27 simulated landscapes under three spatial arrangements and increasing number of patch types/evenness gradient	124
E2. Index values from 27 simulated landscapes with equal patch proportions under three spatial arrangements and increasing number of patch types	125

LIST OF FIGURES

1. A four-category mosaic sampled at a row of points	9
2. The nine physiographic provinces of Alabama	39
3. Distribution of Southern Forest Inventory and Analysis survey plots on three physiographic provinces in Alabama	42
4. Forest Cover type changes through time in the Great Appalachian Valley physiographic province based on Southern Forest Inventory and Analysis survey data	44
5. Forest Cover type changes through time in the Blue Ridge-Talladega Mountain physiographic province based on Southern Forest Inventory and Analysis survey data	45
6. Forest Cover type changes through time in the Piedmont physiographic province based on Southern Forest Inventory and Analysis survey data	46
7. Simulated random landscape maps for 2 to 10 patch types with the base patch type (white) coverage decreasing by 10 percent with each new patch type. The first map has a coverage of 90 percent for the base patch type and 10 percent for the second patch type	51
8. Simulated uniform landscape maps for 2 to 10 patch types with the base patch type (white) coverage decreasing by 10 percent with each new patch type. The first map has a coverage of 90 percent for the base patch type and 10 percent for the second patch type	52
9. Simulated aggregated landscape maps for 2 to 10 patch types with the base patch type (white) coverage decreasing by 10 percent with each new patch type. The first map has a coverage of 90 percent for the base patch type and 10 percent for the second patch type	53
10. Trend lines showing relationship between Γ_1 and the two controlled variables: spatial pattern and number of patch types. There is an increasing gradient of evenness along the x-axis	54
11. Trend lines showing relationship between Γ_2 and the two controlled variables: spatial pattern and number of patch types. There is an increasing gradient of evenness along the x-axis	55

12. Trend lines showing relationship between RC_2 and the two controlled variables: spatial pattern and number of patch types. There is an increasing gradient of evenness along the x -axis	57
13. Simulated random landscape maps for 2 to 10 patch types with all patch types on each map having equal coverage	58
14. Simulated uniform landscape maps for 2 to 10 patch types with all patch types on each map having equal coverage	59
15. Simulated aggregated landscape maps for 2 to 10 patch types with all patch types on each map having equal coverage	60
16. Trend lines showing relationship between Γ_1 and the two controlled variables: spatial pattern and number of patch types. Relative evenness ≈ 1 at each point along the x -axis	61
17. Trend lines showing relationship between Γ_2 and the two controlled variables: spatial pattern and number of patch types. Relative evenness ≈ 1 at each point along the x -axis	62
18. Trend lines showing relationship between RC_2 and the two controlled variables: spatial pattern and number of patch types. Relative evenness ≈ 1 at each point along the x -axis	64
19. Three uniform landscapes; (A) has 4 patch types, (B) has 5 patch types, and (C) has 6 patch types	65
20. Relationship between D_2^* and the two controlled variables: spatial pattern and number of patch types. Note that there is an increasing gradient of evenness along the x -axis	68
21. Simulated (A) random, (B) uniform, and (C) aggregated landscapes used in Monte Carlo investigation of contagion index variance and distribution	70
22. Kernel smoothed replicate histograms of $\hat{\Gamma}_1$ using the (A) random, (B) uniform, and (C) aggregated landscapes of Figure (21)	73
23. Kernel smoothed replicate histograms of $\hat{\Gamma}_2$ using the (A) random, (B) uniform, and (C) aggregated landscapes of Figure (21)	74

ABSTRACT

Studies of spatial patterns of landscapes are useful to quantify human impact, predict wildlife effects, or describe various landscape features. A robust landscape index should quantify two distinct components of landscape diversity: composition and configuration. One category of landscape index is the contagion index.

A generalized measure of contagion is defined as a function of concentration. From this definition two contagion indices, Γ_1 (a new index) and Γ_2 (an entropy formulation), are derived from expected values of geometric random variables. A widely used relative contagion index, RC_2 , is shown to be a scaled version of Γ_2 .

Distributional properties of $\hat{\Gamma}_1$, $\hat{\Gamma}_2$, and $\hat{\Gamma}_{2(scaled)}$ (i.e., \hat{RC}_2) are derived. They are shown to be asymptotically unbiased, consistent, and asymptotically normally distributed. Variance formulas for $\hat{\Gamma}_1$, $\hat{\Gamma}_2$, and $\hat{\Gamma}_{2(scaled)}$ are derived using the delta method. A Monte Carlo study using subseries analysis and replicate histograms, for variance and distribution assessment, was done as a validity check.

Behavior of Γ_1 , Γ_2 , and RC_2 were investigated with simulated random, uniform, and aggregated landscapes. Both Γ_1 and Γ_2 provide acceptable measures of contagion. The index RC_2 is shown to be an index of evenness, and not of contagion. It is demonstrated that relativized contagion indices are mathematically untenable.

As an application, the pattern and changes in forest cover types over the last two decades were analyzed on three landscape level physiographic provinces of the state of Alabama: (i) The Great Appalachian Valley Province, (ii) The Blue Ridge-Talladega Mountain Province, and (iii) The Piedmont Province. The USDA Forest Service conducts

periodic surveys of forest resources nationwide from plots distributed on a 3-mile by 3-mile (4.8-km by 4.8-km) grid randomly established within each county. Using forest inventory and analysis survey data on forest cover types, stratified by physiographic province, the $\hat{\Gamma}_1$ and $\hat{\Gamma}_2$ contagion values and their variances were calculated for each province for the survey years 1972, 1982, and 1990. One-way analysis of variance was used for hypothesis testing of contagion values across time and between provinces. Contagion values were very similar indicating similar processes operating across the physiographic provinces over the last two decades. In comparing $\hat{\Gamma}_1$ and $\hat{\Gamma}_2$, use of $\hat{\Gamma}_1$ in analysis of variance gave a more conservative test of contagion.

INTRODUCTION

The Landscape Scale

Landscape diversity refers to the various ecosystems (including human, e.g., cities, farm country, etc.) within a large area. The structure observed in landscapes result from complex interactions between physical, biological, and social forces. Most landscapes have been altered by human land use, and the resulting landscape mosaic is a mixture of natural and human-managed patches that vary in size, shape, and arrangement (Burgess and Sharpe 1981). This spatial patterning becomes observable at the landscape level (Turner 1989).

Resource professionals have been expanding current management approaches to address landscape-level concerns and issues (Salwasser 1990). With the recognition of such problems as loss of biodiversity, climate change, and ecosystem degradation, a necessary evolution in management scale has come about, based on watersheds or other ecological landscape scale units rather than political boundaries.

Forest landscape management is based on the premise that resource flows as well as biodiversity levels and ecosystem processes are determined by the array and spatial arrangement of forest conditions, i.e., spatial structure, and its change over time. A quantitative basis for measuring spatial structure is a prerequisite to implementing forest landscape management. Without such, structural objectives cannot be established nor can the understanding of spatial dynamics necessary to achieve structural objectives be realized.

Assessing and Comparing Landscape Diversity

The capability to identify and quantify landscape structure in a meaningful way, valuable input for ecosystem management (the new paradigm in USDA Forest Service

policy), has recently gained much attention (Naveh and Lieberman 1994). Further developments are needed in the area of landscape measures of diversity, especially based on ground truth sample data.

In remote sensing the entire scene (area under consideration) is observed at a certain resolution and classified, bit by bit, into various categories. The scene, in raster format, can be displayed on a monitor. Some consider this the ideal situation, having a remotely sensed image where the entire observed scene can be classified. However, in general, tree species cannot be identified on Landsat Thematic Mapper (TM) images or Advanced Very High Resolution Radiometer (AVHRR) images, and so on (Cihlar et al. 1996, Evans and Czaplewski 1996). Reflectance properties are still poorly understood in forest settings. While it is true that broad categories, such as pine, hardwood, and nonforest, are easily classified; and some species, such as baldcypress, can be identified, we still lack the ability to map specific forest cover types from remotely sensed data. The same situation holds with color aerial photographs. Magnussen (1997) found poor agreement ($\approx 47\%$) between photo interpretations of forest cover types and the true ground values from a study of 3317 stands in New Brunswick. But sample-based forest survey data are available for forest cover type mapping. Even when the ability to classify forest cover types on remotely sensed or photo interpreted images is developed, we cannot look back historically except with sample data.

A landscape diversity index is a quantitative metric, a single statistic applied on a broad spatial scale in which two distinct components are confounded: composition and configuration. Composition refers to both the total number of land cover categories or “patch” types and their relative proportions in the landscape, whereas configuration refers

to the spatial pattern of patches in the landscape (Li and Reynolds 1993). One class of landscape index is the contagion index. Contagion, as defined by O'Neill et al. (1988b), measures the extent to which landscape elements are aggregated or clumped. Higher values of contagion generally result from landscapes with a few large, contiguous patches, whereas lower values usually characterize landscapes with many small patches. Also, holding the number of categories more or less constant, contagion values, in general, should decrease as category proportions become more even. To date, only a handful of contagion indices, such as the relative contagion indices of Li and Reynolds (1993), have been proposed, and in no cases have distributional properties been examined.

Many composite landscape indices are used often in the literature, such as indices of dominance, contagion, and patch configuration. These composite indices are good summary indices for description and comparison of landscape structure (Turner 1990, Li and Reynolds 1994). However, single-valued indices have their limitations. For example, offsetting changes in different landscape categories would produce no change in dominance or contagion indices. The caveat here is that single-valued indices, such as contagion, have their uses, but they should not be used in isolation.

Landscape contagion, as a property, has a reverse scale from that of species diversity. A single species community has no diversity while an infinite species community has maximum diversity. The exact opposite is true from a contagion viewpoint. That is, a single category landscape has maximum contagion whereas an infinite category landscape has no contagion. Landscape ecologists have adopted this perspective because large contiguous ecosystems on a landscape are more biologically viable and species rich, in

general, than highly fragmented landscapes. Put another way, the greater the uncertainty in species interspecific encounters the greater the species diversity; but on landscapes, the greater the uncertainty that neighboring areas belong to a specific land class the lower the landscape contagion. In this regard contagion equates to spatial autocorrelation. If categories are not correlated in space (i.e., autocorrelation ≈ 0), then there is greater uncertainty at any point on the identity of surrounding neighbors.

As a statistic, a contagion index is limited if it cannot be used in making comparisons of diversities among different landscapes or the same landscape through time based on sample data. If the land cover categories are the same between two landscapes, then contingency table analysis could be used to test if relative proportions of categories are the same between the landscapes. This would answer the composition question, but not the configuration question. Spatial sampling is typified by systematic sampling at lattice intersection points. As one moves along the transects from point to point, transitions are made from patch type to patch type. The sequences of transitions can be regarded as realizations of a Markov chain. A matrix of transition probabilities can be constructed for each landscape and the null hypothesis that the matrices of transition probabilities are the same can be tested with a log likelihood ratio test (Kullback et al. 1962). However, if the number or kinds of patch types differ among landscapes, how then does one test for differences in landscape diversity? This same problem faced ecologists with species diversity indices. Hutcheson (1970) devised an ANOVA type test for the Shannon species diversity index (H') by investigating the distributional properties and sampling variance of

H' . Moran (1948) gives the sampling variance of a join-count statistic for testing spatial dependence from simple binary classifications of landscapes.

The sampling properties of a landscape index must be known before one can construct an appropriate test. Is the index unbiased? Consistent? Can the variance be computed or approximated? Is the sample distribution normal? A formal investigation is necessary to answer these questions.

Application

According to Wenger (1984), vegetation classification provides better communication among professionals, greater understanding of plant-environment interactions, and a useful aid for land and resource management. Several forest cover type classifications are currently used within the forestry profession, the two most popular in the U.S. are the Society of American Foresters (1980) system and the USDA Forest Service Forest Inventory and Analysis (FIA) system (USDA FS 1972). In most southern states the geologic landforms have been mapped and classified into physiographic provinces. In particular in Mississippi and Alabama digitized maps are readily available for input into a Geographic Information System (GIS). At this time, the Louisiana data have not been digitized. With the digitized maps and the FIA survey data, a unique opportunity exists to apply a contagion index and distribution theory to characterize and compare forest cover type diversity on several landscape scale physiographic provinces.

Objectives

The objectives of this research were to:

- 1) Develop spatial landscape scale diversity indices in the form of a contagion index,

- 2) Study the sampling properties of these new contagion indices,
- 3) Investigate their behavior on simulated landscapes, and
- 4) Demonstrate use of the new indices on USDA Forest Service FIA data on forest cover types for comparing three landscape level physiographic provinces of Alabama, (i) Great Appalachian Valley Province, (ii) Blue Ridge-Talladega Mountain Province, (iii) Piedmont Province.

REVIEW OF LITERATURE

For purposes of discussion, a landscape here refers to some well defined geographically targeted unit, such as a physiographic province, a watershed, or other ecologically meaningful unit.

Concepts

Most communities of sessile organisms can be thought of as mosaics (Pielou 1975). The mosaic structure on a landscape is sometimes clear as in, for example, desert shrub communities, or the mosaic patches may be quite small (e.g., mixed grassland), that they can hardly be described as mosaics in the literal sense; however most researchers still treat them as mosaics (Pielou 1975). When spatial pattern is disregarded, a landscape's diversity is, intuitively, equal to the "uncertainty" of the specific identity of an area picked at random from it. But if we take the spatial pattern into account, and attempt to predict the identity of an area near in space to one that has already been identified, the chance of successful prediction becomes much greater. There is much less uncertainty about the identity of a nearby neighbor than a distant one. This principal is the basis for devising measures of "spatial diversity" (Pielou 1975).

For illustration, consider a vegetation mosaic, an area completely occupied by n nonoverlapping patches of different categories. The categories are the different kinds of vegetation present; a category may consist of a single species, or a recurring group of species that form a recognizably distinct community type (such as a forest cover type). The categories are disbursed about into patches. We shall treat each patch as a continuous finite area. Let the mosaic be sampled at a sequence of N equidistant points along a line transect

laid across it (Figure 1), and let the category or classification in which each point falls be recorded. For example, with $N = 18$ and $n = 4$ land cover categories labelled A, B, C, and D, the observations consist of a list such as: AAA BB C B DDDDD CC BBB D (see Figure 1). From such a list Pielou (1975) devised the following measure of spatial diversity. Let N_{ij} be the observed frequency with which a point in the i th category is succeeded by a point in the j th category. The total number of these is $\sum_i \sum_j N_{ij} = N$; and $N_{ij}/N_i = p_{ij}$ is an estimate of the probability that a point in the i th category will be succeeded by a point in the j th category, where N_i is the total number of sampling points that fall in classification i . Hence, $N_i/N = p_i$ is an estimate of the proportional area of classification i in the mosaic. Pielou measured the mosaic's spatial diversity as:

$$H'_{(1)} = - \sum_i \sum_j p_i p_{ij} \ln(p_{ij}) \quad (1)$$

It was natural for quantitative ecologists to expand on these ideas and devise landscape measures based on a grid of points or pixels, as we will see in the next section.

Landscape Indices

A number of studies have examined habitat fragmentation to quantify human impact, predict wildlife effects, or describe various landscape features (De Cola 1989; Krummel et al. 1987; O'Neill et al. 1988a, 1988b). Most of these have used a binary classification approach to the landscape. For example, if all old-growth habitats on a map are colored black and the other habitats white (whether represented by polygons or pixels), then one can easily look at fragmentation, area/perimeter ratio, or connectivity of the old-growth sites.

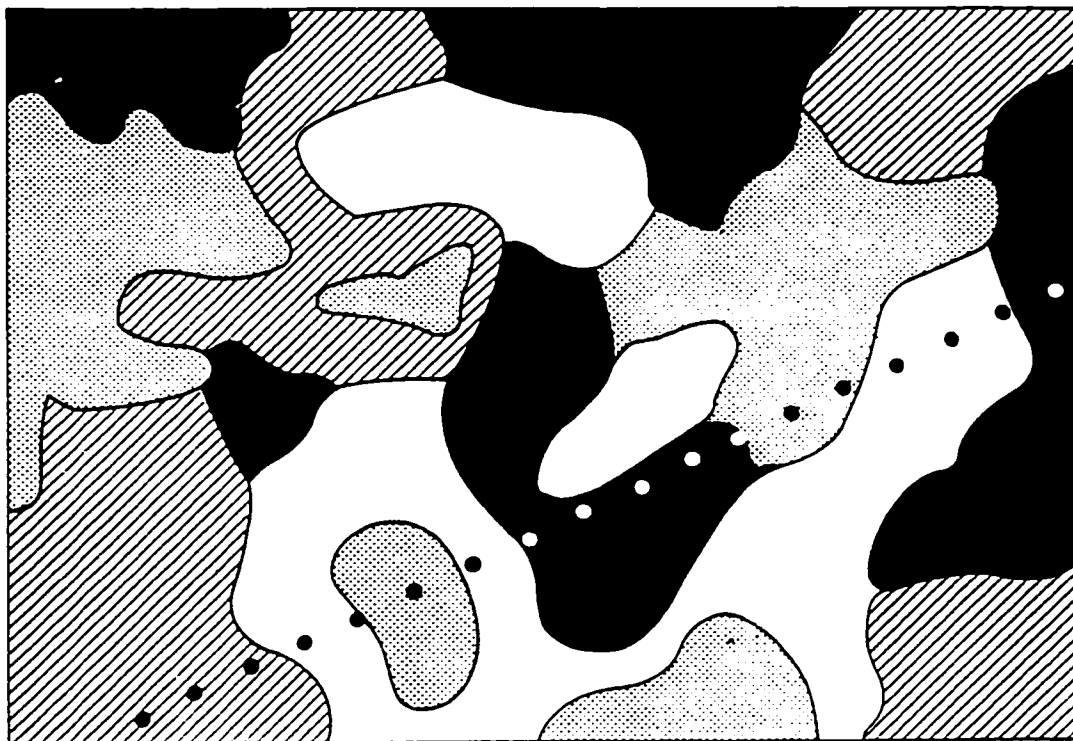


Figure 1. A four-category mosaic sampled at a row of points

One of the early measures to quantify spatial dependence of a spatial binary random variable Z mapped over a regular or irregular grid is the join-count statistic:

$$r = \sum_i \sum_j W_{ij} C_{ij} \quad (2)$$

where W_{ij} is a measure of spatial proximity between the two locations i and j and C_{ij} is a different measure of proximity based on observed values of Z at those locations (Davis 1993). A binary map coded as black (B) and white (W) ($z_i = 1$ for black, $z_i = 0$ for white) can be described by the frequencies of BB, BW, and WW joins among adjacent cells out of a total of J joins. Let $W_{ij} = 1$ if locations i and j are adjacent and 0 otherwise. For gridded data, adjacency may be defined as sharing a common side (“rook’s” definition, from chess), as sharing a corner (“bishop’s” definition), or as either a side or corner (“queen’s” definition). The standard used is rook’s rule (Gonzalez and Wintz 1987). If in Equation (2) $C_{ij} = (z_i - z_j)^2$, then $r/2$ is the number of times adjacent locations have different identities. If $C_{ij} = z_i z_j$, then $r/2$ is the number of BB joins. The number of WW joins is computed as $J - BB - BW$. Moran (1948) gave the mean and variance of the join-count statistic so formal hypothesis tests could be conducted:

$$E[r] = S_0 T_0 / n(n-1) \quad (3a)$$

$$\text{var}(r) = \frac{S_1 T_1}{2n(n-1)} + \frac{(S_2 - 2S_1)(T_2 - 2T_1)}{4n(n-1)(n-2)} + \frac{(S_0^2 + S_1 - S_2)(T_0^2 + T_1 - T_2)}{n(n-1)(n-2)(n-3)} - E^2[r] \quad (3b)$$

where n is number of locations; $S_0 = \sum_i \sum_j W_{ij} (i \neq j)$; $S_1 = \frac{1}{2} \sum_i \sum_j (W_{ij} + W_{ji})^2 (i \neq j)$; $S_2 = \sum_i (W_{i0} + W_{0i})^2$; $W_{i0} = \sum_j W_{ij}$; $W_{0i} = \sum_j W_{ji}$; and T_0, T_1, T_2 are calculated by substituting C_{ij} for W_{ij} in formulas for S_0, S_1 , and S_2 , respectively. For an excellent review of the early work in spatial statistics see Ripley (1981).

The binary framework is not adequate for dealing with the problem of landscape diversity because one must consider the complexity of mixes of different types, not just single types at a time. Turner et al. (1995) considered knowledge of the spatial arrangement of habitats across the landscape as essential to understanding the ecology of the area and for effective management. They felt simple comparisons, such as suitable versus unsuitable, were not sufficient.

Some indices have been developed that evaluate multiple habitat types simultaneously. O'Neill et al. (1988b) developed two measures based on information theory. I would like to point out that the genesis of the information index (i.e., Shannon diversity, H') is the log likelihood function of the multinomial distribution. Their first index is a measure of dominance:

$$D_1 = \ln(n) + \sum_{i=1}^n p_i \ln(p_i) \quad (4)$$

where p_i is the proportion of grid cells (pixels) on the landscape in land cover i and n is the total number of land cover categories (patch types) in a particular landscape unit (physiographic province, watershed, etc.). When land categories are well defined (e.g., agriculture, urban, forest) D_1 works well. O'Neill et al.'s second index quantifies the extent to which different types are intermingled. It is given by:

$$D_2 = 2n \ln(n) + \sum_{i=1}^n \sum_{j=1}^n p_{ij} \ln(p_{ij}) \quad (5)$$

where p_{ij} is the probability that a grid point of land cover i is found adjacent (rook's rule) to a grid point of land cover j . Again, this index assumes that the n classifications or types are distinct.

Li and Reynolds (1993) showed that D_2 is insensitive to changes in spatial pattern. This is because of an error, which they identified, in the formulation of the index. Li and Reynolds defined relative contagion (RC) as:

$$RC = 1 - EE/EE_{\max} \quad (6)$$

where EE denotes the entropy value. The measure of entropy for categorical data is well known, being $-\sum p_i \ln(p_i)$, and was derived by Shannon (Shannon and Weaver 1949). Based on Equation (6), Li and Reynolds gave two new indices. The first is:

$$RC_1 = 1 + \frac{\sum_{i=1}^n \sum_{j=1}^n p'_{ij} \ln(p'_{ij})}{n \ln(n)} \quad (7)$$

where $p'_{ij} = p_{ji} = N_{ij}/N_i$, p_{ji} is the conditional probability, N_{ij} is the number of joins between pixels of patch types i and j (rook's rule), and N_i is the total number of joins between pixels of patch type i and all patch types (including patch i itself). With this definition of p'_{ij} they proved $EE_{\max} = n \ln(n)$. Their second index is:

$$RC_2 = 1 + \frac{\sum_{i=1}^n \sum_{j=1}^n p_{ij} \ln(p_{ij})}{2 \ln(n)} \quad (8)$$

where $p_{ij} = p_i p_{j/i}$ and p_i is the probability that a randomly chosen pixel belongs to patch type i (estimated by the proportion of patch type i). With this definition of p_{ij} Li and Reynolds proved $EE_{\max} = 2\ln(n)$. The error in D_2 was the use of $2n\ln(n)$ instead of the correct EE_{\max} of $2\ln(n)$. Thus D_2 has a lower limit of 0 and no upper limit, but RC_1 and RC_2 range from 0 to 1.

Li and Reynolds (1993) used a series of simulated landscapes to investigate the behavior of their indices RC_1 and RC_2 . They plotted values of their indices against 3 spatial configurations (random, uniform, and aggregated) and increasing numbers of patch types (from 2 to 10) with a simultaneous change in patch type proportions going from one type dominating towards even category proportions. They found that RC_1 varied with different spatial configurations but did not change much with the composition gradient. In contrast, RC_2 distinguished the 3 different spatial configurations and was sensitive to the composition gradient.

Li and Reynolds (1996) employed RC_2 in a study of rangeland degradation in the Jornada Basin of southern New Mexico. Using a simulation experiment they modeled annual changes in patch types with a Markov transition model and used the contagion index to quantify the changing spatial configuration. The Markov transition probabilities came from empirical observations made in the Jornada. The whole process was driven by drought events and grazing intensities imposed during the simulation runs. This study shows contagion has many potential uses, and does not have to stand alone, but can be incorporated with other approaches.

Riitters et al. (1996) viewed contagion as the degree to which mapped attributes are clumped into patches of the same attribute class. They examined what happens to the maximum possible entropy value for different ways of tabulating attribute (patch type) adjacencies--with and without preserving the order of pixels in pairs. Each pixel pair (i,j) has an attribute adjacency type A_q ($q = 1 \dots n_a$) corresponding to the attribute class values (c_i, c_j) of the pixels. For example, if the map has $n = 2$ attribute classes and order is preserved, then there are four types of attribute adjacencies ($A_1 = \{1,1\}$, $A_2 = \{1,2\}$, $A_3 = \{2,1\}$, $A_4 = \{2,2\}$). When pixel order is not preserved, $n_a = 3$, and the adjacency types are $\{1,1\}$, $\{1,2\}$, and $\{2,2\}$. They showed when pixel order is preserved, the maximum entropy value is $2\ln(n)$ and the index RC_2 is obtained. When pixel order is not preserved, $EE_{\max} = \ln(n^2 + n) - \ln(2)$ and a new index results:

$$RC_{unordered} = 1 + \frac{\sum_{q=1}^{n_a} p_q \ln(p_q)}{\ln(n^2 + n) - \ln(2)} \quad (9)$$

where $p_q = f_q/n_p$, f_q is the frequency of pixel pairs that are of attribute adjacency type q , and n_p is the total number of pixel pairs in the map. Riitters et al. (1996) argued that Equation (9) is consistent with the motivation of an entropy calculation which requires that each "state" be distinguishable. They also defined a new rule for counting joins. They used 2 sides of each pixel (the right and bottom sides, though any 2 sides could be used), instead of all four sides as in rook's rule. They recognized that counting order is arbitrary under the 2-sided rule, so that rook's rule is attractive for the reason that it obscures any ordering.

The interspersion and juxtaposition index (*IJI*) (McGarigal and Marks 1995) measures the distribution of adjacencies among patch types. *IJI* ranges in value from 0 to 100, where the lower range represents a situation of clumped patch types (all patch types not equally adjacent to all other patch types), and the upper range represents maximum interspersion and juxtaposition. The index is computed as follows:

$$IJI (\%) = \frac{-\sum_{i=1}^n \sum_{k=i+1}^n \left(\frac{e_{ik}}{E} \times \ln \frac{e_{ik}}{E} \right)}{\ln \left(\frac{1}{2} [n(n-1)] \right)} \times 100 \quad (10)$$

where e_{ik} is the total length of edge in the landscape between patch types i and k , E is the total length of edge of the landscape, and n is the number of patch types present in the landscape. This index is similar to the contagion index of Li and Reynolds (1993), except that the interpretation is reversed. In contagion, the preferable condition, that of a few large contiguous patches, is given a high value, whereas many small interspersed patches, an undesirable condition generally speaking, is given a small value. The exact opposite is true with *IJI*.

Bettinger et al. (1996) chose *IJI* for use in a study on the effects of changing grid cell size on landscape index values. They recognized RC_2 as a similar metric but used *IJI* because it was part of the FRAGSTATS software package (McGarigal and Marks 1995). They used a vegetation coverage scene from eastern Oregon rasterized at five different grid cell sizes. They found that going from 3 to 6 and then to 10 m resolution produced little change. However, when grid cell size increased to 20 and 30 m, *IJI* decreased.

Loehle and Wein (1992) have attempted to derive an index that incorporates the number of different habitat types, their spatial interpenetration, and the extent to which the types are different from one another. They assume that all habitat types can be ordinated along a single axis (e.g., wettest to driest) by some method such as principal component analysis or detrended correspondence analysis. Next, they assign to each grid cell or pixel an “elevation” corresponding to its location along the ordination axis. Then the ordination axis is shifted to give positive values, with a lowest elevation = 0. This procedure produces a pixel-based “elevation” map which they then analyze for roughness and patchiness using a fractal measure. They used an information dimension d_I (Farmer et al. 1983) for analysis:

$$d_I = \lim_{\epsilon \rightarrow 0} \frac{I(\epsilon)}{\ln(1/\epsilon)}, \text{ and } I(\epsilon) = \sum_{i=1}^{K(\epsilon)} p_i \ln(1/p_i) \quad (11)$$

where p_i is the probability of occurrence of the function within the i th map square, and $K(\epsilon)$ is the number of map squares having sides of length ϵ pixels. For a totally uniform map, $d_I = 0$. Values for d_I below 1.0 indicate widely scattered patches of different habitat. A value of approximately 1 indicates linear features (e.g., a river in a uniform landscape). Patches that differ notably from their neighbors will cause values to be in the range $2 < d_I < 3$.

An index frequently used with TM and AVHRR satellite data for vegetation recognition, that is, gross land cover classification, is the normalized difference vegetation index (*NDVI*):

$$NDVI = \frac{NIR - VIS}{NIR + VIS} \times 100 \quad (12)$$

where NIR is the reflectance values for the near-infrared channel (channel 2 in AVHRR) and VIS is the reflectance values for the visible channel (channel 1 in AVHRR) (Cihlar et al. 1996, Evans and Czaplewski 1996). $NDVI$ data are used as input to classification algorithms to identify limited vegetative cover types. The cover types generally are agriculture, grassland, broadleaf forest, coniferous forest, and mixed forest. It is currently not possible, with some exceptions, to identify specific crops, specific forest types, etc. This index is concerned with categorizing patch types. It is used to classify pixels in a scene which can then be analyzed with other indices to examine fragmentation, spatial configuration, and so on. $NDVI$ is reviewed here to show the common procedure for classifying pixels from vegetative scenes and to establish the current limitations of land cover classification. Exacting classification comes from ground data, and FIA units have sample-based forest cover type data for all states in the U.S.

Of the indices reviewed, RC_2 is both 1) sensitive to composition/configuration, and 2) straightforward to apply and easy to interpret. The IJI index is sensitive to composition/configuration, but length of edge in a sampling context would be difficult to determine. Also, Baskent and Jordan (1995), in reviewing measures of spatial structure of forest landscapes, considered contagion of high utility and put little emphasis on IJI . While the information dimension d_I is sensitive to composition/configuration, it is tedious to apply, requiring the use of a sophisticated statistical procedure. The other indices, D_1 , D_2 and RC_1 , are not sensitive to composition/configuration or behave illogically, though they are easy to apply. $RC_{unordered}$ is basically the same as RC_2 , giving slightly higher values of

contagion under similar circumstances, according to Riitters et al. (1996). The index RC_2 , then, seems a reasonable yardstick to compare new indices against.

Variance Computation from Distribution Theory

For n defined as the number of land cover categories, a contagion index is a scalar statistic that is a nonlinear function of the n^2 elements of a $n \times n$ contingency table, where p_{ij} is the ij th element of the contingency table. A common approach for establishing the asymptotic distribution and variance of a scalar statistic that is a function of k variables is the delta method (Rao 1965, pp. 321-322; Bishop et al. 1975, pp. 492-497).

Let \mathbf{P} be a k -dimensional statistic (P_1, \dots, P_k) and let T be the total number of observations. Further let g be a function of k variables which is totally differentiable. If the asymptotic distribution of $T^{1/2}(P_1 - \theta_1), \dots, T^{1/2}(P_k - \theta_k)$ is k -variate normal with mean zero and dispersion matrix $\Sigma = (\sigma_{ij})$, then the asymptotic distribution of

$$\sqrt{T}\mu = \sqrt{T}(g(P_1, \dots, P_k) - g(\theta_1, \dots, \theta_k)) \quad (13)$$

is normal with mean zero and variance

$$v(\theta) = \sum \sum \sigma_{ij} \frac{\partial g}{\partial \theta_i} \frac{\partial g}{\partial \theta_j} \quad (14)$$

The formal method of determining the variance of $g(P_1, \dots, P_k)$ is as follows. Take the total differential of g :

$$dg = \frac{\partial g}{\partial \theta_1} dP_1 + \dots + \frac{\partial g}{\partial \theta_k} dP_k \quad (15)$$

where dP stands for $(P_1 - \theta_1)$ etc. Now compute the variance, assuming that the variances and covariances of $(P_i - \theta_i)$ exist:

$$v(dg) = \sum_r \sum_s \frac{\partial g}{\partial \theta_r} \frac{\partial g}{\partial \theta_s} \text{cov}(P_r, P_s) \quad (16)$$

Variance Approximation from Nonparametric Theory

Sometimes a check of theoretical derivations is desired to assess validity of results. In such cases nonparametric theory coupled with a simulation study provides a valid check. Sometimes a target parameter θ is estimated by a complicated scalar statistic $R(\cdot)$ so that theoretical derivations are analytically intractable. In such cases nonparametric estimation of confidence intervals on θ is the only recourse.

Suppose Z_1, \dots, Z_T is a series of T observations from a stationary sequence with cumulative distribution function F and let $R(\mathbf{Z}; F)$ denote a random variable (test statistic) of interest. Nonparametric techniques are available to obtain properties of the sampling distribution of $R(\mathbf{Z}; F)$ without having to assume a distributional form for F . When the data are independent and identically distributed (iid), that is, the Z_i 's are iid random variables, the jackknife (Gray and Schucany 1972) and bootstrap (Efron and Tibshirani 1993) have proved their worth. Today there is a growing body of literature on resampling methods for dependent data using "subseries". In the subseries approach the observed series is divided into blocks and the blocks are taken as replicates of the original data structure.

Consider a regularly spaced stationary time series $\{Z(t): t = 1, 2, \dots\}$, of which there are T observations $\mathbf{Z} \equiv (Z(1), \dots, Z(T))'$. Define $\mathbf{Z}_i^m \equiv (Z(i+1), \dots, Z(i+m))'$ to be a subseries of values of length m . Then $\{\mathbf{Z}_{jm}^m: j = 0, \dots, [T/m]-1\}$ denotes a sequence of

subseries, where $[x]$ is the integer part of x . If the statistic $\hat{\theta}(\mathbf{Z})$ is a function of $\{\mathbf{Z}_k^m : 0 \leq k \leq T - m\}$, then a subseries approach can be defined (Carlstein 1986a, 1988) by selecting nonoverlapping blocks. Carlstein computed an empirical variance from the “replicates” as:

$$\sum_{j=0}^{[T/m]-1} \frac{\{\hat{\theta}(\mathbf{Z}_{j,m}^m) - \bar{\theta}(\mathbf{Z})\}^2}{[T/m]}, \text{ where } \bar{\theta}(\mathbf{Z}) \equiv \sum_{j=0}^{[T/m]-1} \frac{\hat{\theta}(\mathbf{Z}_{j,m}^m)}{[T/m]} \quad (17)$$

and gave conditions under which it converges to $\lim_{T \rightarrow \infty} \text{var}(T^{1/2} \hat{\theta}(\mathbf{Z}))$. It is necessary that m be chosen such that the subsamples preserve the dependence structure in $\{Z(t)\}$. Carlstein (1986b) also specified necessary and sufficient conditions under which a general statistic $\hat{\theta}(\mathbf{Z})$ has asymptotic normality without specifying the dependence model giving rise to $\{Z(t)\}$, and without specifying the marginal distribution of $Z(t)$.

Sherman and Carlstein (1994) generalized Carlstein’s (1986a, 1988) approach to the spatial setting. For data $\mathbf{Z} \equiv (Z(\mathbf{s}_1), \dots, Z(\mathbf{s}_T))'$ from a stationary lattice process defined on $D \subset \mathbb{R}^d$, divide D into mutually exclusive and exhaustive congruent subregions D_1, \dots, D_K . Let the corresponding sublattice data be $\mathbf{Z}_{D_1}, \dots, \mathbf{Z}_{D_K}$ from which statistics $\hat{\theta}(\mathbf{Z}_{D_k})$, $k = 1, \dots, K$, can be computed. Sherman and Carlstein showed that

$$\hat{v}_T = \sum_{k=1}^K \frac{\{\hat{\theta}(\mathbf{Z}_{D_k}) - \bar{\theta}(\mathbf{Z})\}^2}{K}, \text{ where } \bar{\theta}(\mathbf{Z}) \equiv \sum_{k=1}^K \frac{\hat{\theta}(\mathbf{Z}_{D_k})}{K} \quad (18)$$

is a consistent estimator of $\text{var}(T^{1/2} \hat{\theta}(\mathbf{Z}))$. They also showed asymptotic normality under mild conditions on $\hat{\theta}(\mathbf{Z})$ and mixing conditions on the strength of spatial dependence. This type of estimator has also been used by Possolo (1991) and Politis and Romano (1993,

1994). As in the time series case, it is necessary that subregions be chosen such that they preserve the spatial dependence structure in D .

Sherman and Carlstein (1996) have introduced a simple diagnostic tool, the “replicate histogram”, for describing the sampling distribution of a general statistic. The method, completely sample based, requires no theoretical analysis by the user. The goal is to obtain diagnostic information about the cumulative distribution function F . The replicate histogram alerts the user to nonnormal sampling distributions and also indicates the type of departure from normality such as peakedness or skewness.

Let $D_l^k, k = 1, \dots, K$ denote overlapping subregions of D , where l determines the common size of each subregion and K is the number of all possible subregions of size l . To examine the features of the unknown F , Sherman and Carlstein (1996) proposed constructing the empirical distribution of the subseries replicates

$$\hat{F}(z) \equiv \frac{\sum_{k=1}^K \mathbb{I}\{\hat{\theta}(\mathbf{Z}_{D_l^k}) \leq z\}}{K}, \quad z \in \mathbb{R} \quad (19)$$

and plotting the corresponding kernel smoothed histogram. Sherman and Carlstein established the large-sample validity of the replicate histogram via strong consistency results. The behavior of the replicate histogram is, of course, influenced by variation in the choice of l .

DERIVING NEW CONTAGION INDICES

Li and Reynolds (1993) defined a general form for relative contagion (Equation (6)), and derived two entropy-based formulations for relative contagion based on two definitions of p_{ij} , the probability that a quadrat or pixel of land cover i is found adjacent to a pixel of land cover j . They showed that p_{ij} defined as $p_i p_{j/i}$ (proportion of patch type i times the conditional probability) led to an index sensitive to both composition and configuration. I will retain this definition of p_{ij} .

In general, formulas for quantifying diversity have been motivated by information-theoretic axioms (i.e., the principle of entropy) (Rényi 1961, Hill 1973) or arguments of mean proportional abundance and average rarity (i.e., expected value operations) (Hill 1973, Peet 1974, Patil and Taillie 1982). One can view the adjacency of patch type i and j as a general problem in waiting times (Feller 1968) for the encounter of state ij , which follows a geometric distribution. I will show that by defining contagion as a generalized function and inserting expected values of random variables based on the encounter of state ij , new and logical indices result.

Contagion Generalized

In a *contagious* landscape, as described by O'Neill et al. (1988b) and Riitters et al. (1996), the typical patch type is relatively concentrated. Therefore I view contagion as a function of concentration. Denoting the concentration function of p_{ij} as $C(p_{ij})$, a generalized measure of contagion on landscape L is given by:

$$\Gamma(L) = \phi - \sum \sum p_{ij} C(p_{ij}) \quad (20)$$

where Γ is the contagion index associated with the measure of concentration C and ϕ is any real constant. Because the meaning of contagion is the inverse of the meaning of species diversity, it is necessary to subtract the quantity $\sum \sum p_{ij} C(p_{ij})$ from some constant ϕ to reverse the scale and provide a contagion formulation. Li and Reynolds (1993) used $\phi = 1$ in formulating relative contagion, defined by Equation (6). This caused their indices RC_1 and RC_2 to have the range 0 to 1. The parameter ϕ can be considered a location parameter.

Geometric Distribution

Recall that p_{ij} is the probability that two randomly chosen adjacent pixels belong to type i and j out of n patch types. Define $X + 1$ as the number of random picks of adjacent pixels up to and including the first encounter of state ij . This scheme is a general problem in waiting times. Under this scheme X has a geometric distribution:

$$P(X = x | p_{ij}) = p_{ij} (1 - p_{ij})^x, \quad x = 0, 1, 2, \dots \quad (21)$$

If patch type i (or j) dominates a landscape, for $i \neq j$ a large value of X would be expected. The ratio $X/(X + 1)$ provides a reasonable measure of concentration, varying from $1/2$ (low concentration) to 1 (high concentration). For the situation $i = j$ the ratio $X/(X + 1)$ still provides a reasonable measure of concentration, the meaning is simply reversed. That is, if patch type i dominates the landscape, a small value of X would be expected. While it may seem incongruous to mix the different meanings, recall that correlation is a mix of variance and covariance, each of which has different meanings, but provides a logical statistic. Now the ratio $X/(X + 1)$ is a random variable and the average of such a ratio can be constructed in a number of ways; each gives rise to a different index.

New Contagion Index Γ_1

Let us assume that $C(p_{ij})$ is represented by $C(p_{ij}) = E[X/(X+1) | p_{ij}]$. Since $E[X/(X+1) | p_{ij}] = 1 + p_{ij} \ln(p_{ij})/(1 - p_{ij})$ (see Appendix A for derivations), using this result in Equation (20) with $\phi = 1$ gives:

$$\Gamma_1 = 1 - \sum_{i=1}^n \sum_{j=1}^n p_{ij} [1 + p_{ij} \ln(p_{ij})/(1 - p_{ij})] \quad (22)$$

Because $\sum_{i=1}^n \sum_{j=1}^n p_{ij} = 1$ by definition, the final form of the new contagion index is:

$$\Gamma_1 = \sum_{i=1}^n \sum_{j=1}^n \frac{p_{ij}^2 \ln(p_{ij})}{p_{ij} - 1} \quad (23)$$

This index is bounded between 0 and 1. For a proof see Appendix A.

New Contagion Index Γ_2

Let us assume that $C(p_{ij})$ is represented by $C(p_{ij}) = E[X | p_{ij}] \times E[1/(X+1) | p_{ij}]$. Since $E[X | p_{ij}] = (1 - p_{ij})/p_{ij}$ and $E[1/(X+1) | p_{ij}] = -p_{ij} \ln(p_{ij})/(1 - p_{ij})$ (see Appendix A for derivations), $C(p_{ij}) = -\ln(p_{ij})$. Use of this result in Equation (20) with $\phi = 0$ gives:

$$\Gamma_2 = \sum_{i=1}^n \sum_{j=1}^n p_{ij} \ln(p_{ij}) \quad (24)$$

The quantity $\sum \sum p_{ij} \ln(p_{ij})$ is the entropy information value (Shannon and Weaver 1949). It has already been established that by defining p_{ij} as the product $p_i p_{j/i}$, EE_{\max} is $2\ln(n)$ (see Equation (8)). Therefore Γ_2 is bounded between $-2\ln(n)$ and 0. The index RC_2 is readily obtained from Γ_2 by dividing by EE_{\max} and adding 1 which rescales to the interval $[0, 1]$, thus $\Gamma_{2(\text{scaled})} = RC_2$.

SAMPLING PROPERTIES OF $\hat{\Gamma}_1$, $\hat{\Gamma}_2$, AND $\hat{\Gamma}_{2(\text{scaled})}$

In this section let Γ simultaneously represent Γ_1 , Γ_2 , and $\Gamma_{2(\text{scaled})}$ (i.e., RC_2). To obtain an estimate of the contagion

$$\Gamma(p_{11}, p_{12}, \dots, p_{nn}) \quad (25)$$

on the basis of a given sampling lattice, the unknown *a priori* probabilities, the p_{ij} 's, in Equation (25) are replaced by estimated \hat{p}_{ij} 's. In this connection, then, the properties of the random variable

$$\hat{\Gamma} = \Gamma(\hat{p}_{11}, \hat{p}_{12}, \dots, \hat{p}_{nn}) \quad (26)$$

are what we wish to determine.

The *O*, *o* Notation

A commonly used convention for keeping track of the 'order of magnitude' of an approximation is the "big *O*, little *o*" notation. The sequence $\{b_n\}$ is *at most of order* n^λ , denoted $O(n^\lambda)$, if and only if for some real number Δ , $0 < \Delta < \infty$, there exists a finite integer N such that for all $n \geq N$, $|n^{-\lambda}b_n| < \Delta$. The sequence $\{b_n\}$ is *less than order* n^λ , denoted $o(n^\lambda)$, if for any $\epsilon > 0$, there exists a finite integer N such that for all $n \geq N$, $|n^{-\lambda}b_n| < \epsilon$. For an excellent review of order of magnitude see White (1984, pp. 14-16). The idea behind these two definitions is the comparison of the approximate size or *order of magnitude* of $\{b_n\}$ to that of n^λ . The interpretation of $\{b_n\} = O(n^\lambda)$ is that $\{b_n\}$ is of roughly the same size or order of magnitude as n^λ . The interpretation of $\{b_n\} = o(n^\lambda)$ is that $\{b_n\}$ is of a smaller order of magnitude than is n^λ . As a final note, it is often necessary to add together several *O* and/or *o* expressions to obtain a single order-of-magnitude term.

The rule is that the order of magnitude of a sum is the largest order of magnitude of the summands.

Bias

To estimate the first moment of (26) it is necessary to expand the function in a Taylor series about the point (p_{11}, \dots, p_{nn}) and take the expected value. We can write the expansion as follows

$$\begin{aligned} \Gamma(\hat{p}_{11}, \dots, \hat{p}_{nn}) = & \Gamma(p_{11}, \dots, p_{nn}) + \sum \Gamma'(p_y)(\hat{p}_y - p_y) + \frac{1}{2} \sum \Gamma''(p_y)(\hat{p}_y - p_y)^2 \\ & + \frac{1}{6} \sum \Gamma'''(p_y)(\hat{p}_y - p_y)^3 + \frac{1}{24} \sum \Gamma^{(4)}(p_y)(\hat{p}_y - p_y)^4 + \dots \end{aligned} \quad (27)$$

The expectation of $\Gamma(\hat{p}_{11}, \dots, \hat{p}_{nn})$ involves the central moments of the random variables \hat{p}_{ij} ($i, j = 1, \dots, n$). The multinomial distribution arises in categorical data analysis. The genesis of the multinomial distribution stems from T independent trials, where each trial can result in only one of n mutually exclusive events. See Ratnaparkhi (1985) for background on the multinomial distribution. Let T be the total number of lattice plots and let $\varepsilon_{ij} = \hat{p}_{ij} - p_{ij}$. For the multinomial distribution, the moment generating function is (Shenton and Hutcheson 1969):

$$E[e^{\alpha \varepsilon_{ij}}] = \left(E[e^{\alpha(x_{ij} \cdot T - p_{ij} \cdot T)}] \right)^T \quad (28)$$

where x_{ij} takes the value 1 with probability p_{ij} , and 0 with probability $q_{ij} = 1 - p_{ij}$. Using Equation (28), the following values are obtained:

$$\begin{aligned}
E[\epsilon_{ij}] &= 0, \quad E[\epsilon_{ij}^2] = p_{ij}(1 - p_{ij})/T, \\
E[\epsilon_{ij}^3] &= (2p_{ij}^3 - 3p_{ij}^2 + p_{ij})/T^2, \\
E[\epsilon_{ij}^4] &= O(T^{-2}), \quad E[\epsilon_{ij}^5] = O(T^{-3})
\end{aligned} \tag{29}$$

By substituting the appropriate expressions from (29) into (27) and simplifying, we obtain the first moment or mean of the random variable $\hat{\Gamma}$:

$$\begin{aligned}
E[\hat{\Gamma}] &= \Gamma(p_{11}, \dots, p_{nn}) + o(T^{-1}) + o(T^{-2}) + o(T^{-2}) + o(T^{-3}) + \dots \\
&= \Gamma + o(T^{-1})
\end{aligned} \tag{30}$$

From (30) we can deduce that $\hat{\Gamma}_1$, $\hat{\Gamma}_2$, and $\hat{\Gamma}_{2(\text{scaled})}$ are all biased, but for any reasonable size T the bias is very small, and in fact is less than T^{-1} .

Variance

The estimated $\hat{\Gamma}$ equals the true Γ plus a random error ϵ , i.e. $\hat{\Gamma} = \Gamma + \epsilon$. By definition, $E[\epsilon^2] = E[(\hat{\Gamma} - \Gamma)^2]$, and the variance of $\hat{\Gamma}$ is $E[\epsilon^2] - E^2[\epsilon]$. From (30) we know that $E[\epsilon] = o(T^{-1})$, consequently $E^2[\epsilon] = o(T^{-2})$, which for practical purposes is negligible. Therefore, the variance of $\hat{\Gamma}$ is:

$$\text{var}(\hat{\Gamma}) = E[\epsilon^2] \tag{31}$$

The delta method uses the first order multivariate Taylor series expansion to produce an estimated variance $\hat{\text{var}}(\hat{\Gamma})$. Let $\epsilon_{ij} = (\hat{p}_{ij} - p_{ij})$, and $(\partial\Gamma/\partial p_{ij})|_{p_{ij}=\hat{p}_{ij}}$ be the partial

derivative of Γ with respect to p_{ij} evaluated at $p_{ij} = \hat{p}_{ij}$. The first order multivariate Taylor series expansion of $\hat{\Gamma}$ is:

$$\begin{aligned}\hat{\Gamma} \approx & \Gamma + \epsilon_{11} \left(\frac{\partial \Gamma}{\partial p_{11}} \right) \Big|_{p_{11} = \hat{p}_{11}} + \dots + \epsilon_{1n} \left(\frac{\partial \Gamma}{\partial p_{1n}} \right) \Big|_{p_{1n} = \hat{p}_{1n}} \\ & + \epsilon_{21} \left(\frac{\partial \Gamma}{\partial p_{21}} \right) \Big|_{p_{21} = \hat{p}_{21}} + \dots + \epsilon_{2n} \left(\frac{\partial \Gamma}{\partial p_{2n}} \right) \Big|_{p_{2n} = \hat{p}_{2n}} + \dots \\ & + \epsilon_{n1} \left(\frac{\partial \Gamma}{\partial p_{n1}} \right) \Big|_{p_{n1} = \hat{p}_{n1}} + \dots + \epsilon_{nn} \left(\frac{\partial \Gamma}{\partial p_{nn}} \right) \Big|_{p_{nn} = \hat{p}_{nn}}\end{aligned}\quad (32)$$

The Taylor series expansion in Equation (32) provides the following linear approximation:

$$\epsilon = (\hat{\Gamma} - \Gamma) \approx \sum_{i=1}^n \sum_{j=1}^n \epsilon_{ij} \left(\frac{\partial \Gamma}{\partial p_{ij}} \right) \Big|_{p_{ij} = \hat{p}_{ij}} \quad (33)$$

The squared random error approximately equals ϵ^2 from Equation (33):

$$\begin{aligned}\epsilon^2 \approx & \left[\sum_{i=1}^n \sum_{j=1}^n \epsilon_{ij} \left(\frac{\partial \Gamma}{\partial p_{ij}} \right) \Big|_{p_{ij} = \hat{p}_{ij}} \right] \left[\sum_{k=1}^n \sum_{l=1}^n \epsilon_{kl} \left(\frac{\partial \Gamma}{\partial p_{kl}} \right) \Big|_{p_{kl} = \hat{p}_{kl}} \right] \\ \approx & \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \epsilon_{ij} \epsilon_{kl} \left(\frac{\partial \Gamma}{\partial p_{kl}} \right) \Big|_{p_{kl} = \hat{p}_{kl}} \left(\frac{\partial \Gamma}{\partial p_{ij}} \right) \Big|_{p_{ij} = \hat{p}_{ij}}\end{aligned}\quad (34)$$

From Equations (31) and (34), the $\text{var}(\hat{\Gamma})$ is approximated as:

$$\text{var}(\hat{\Gamma}) \approx \sum_{i=1}^n \sum_{j=1}^n \left(\frac{\partial \Gamma}{\partial p_{ij}} \right) \Big|_{p_{ij} = \hat{p}_{ij}} E[\epsilon_{ij} \epsilon_{kl}] \sum_{k=1}^n \sum_{l=1}^n \left(\frac{\partial \Gamma}{\partial p_{kl}} \right) \Big|_{p_{kl} = \hat{p}_{kl}} \quad (35)$$

The covariances $E[\epsilon_{ij}\epsilon_{kl}]$ in Equation (35) can be determined using Equation (28).

The values are as follows:

$$\begin{aligned} \text{cov}(\hat{p}_{ij}\hat{p}_{ij}) &= \text{var}(\hat{p}_{ij}) = E[\epsilon_{ij}^2] - E^2[\epsilon_{ij}] = \frac{\hat{p}_{ij}(1 - \hat{p}_{ij})}{T} \\ \text{cov}(\hat{p}_{ij}\hat{p}_{kl|kl \neq ij}) &= E[\epsilon_{ij}\epsilon_{kl|kl \neq ij}] - E[\epsilon_{ij}]E[\epsilon_{kl|kl \neq ij}] = -\frac{\hat{p}_{ij}\hat{p}_{kl}}{T} \end{aligned} \quad (36)$$

Replacing the expectation in Equation (35) with the values from Equation (36) gives:

$$\begin{aligned} \text{var}(\hat{\Gamma}) &\approx \left\{ \sum_{i=1}^n \sum_{j=1}^n \left[\left(\frac{\partial \Gamma}{\partial p_{ij}} \right) \Big|_{p_{ij}=\hat{p}_{ij}} \right]^2 \frac{\hat{p}_{ij}(1 - \hat{p}_{ij})}{T} \right\} \\ &+ \left\{ \sum_{i=1}^n \sum_{j=1}^n \left[\left(\frac{\partial \Gamma}{\partial p_{ij}} \right) \Big|_{p_{ij}=\hat{p}_{ij}} \right] \sum_{\substack{k=1 \\ kl \neq ij}}^n \sum_{l=1}^n \left[\left(\frac{\partial \Gamma}{\partial p_{kl}} \right) \Big|_{p_{kl}=\hat{p}_{kl}} \right] \frac{-\hat{p}_{ij}\hat{p}_{kl}}{T} \right\} \end{aligned} \quad (37)$$

After algebraic manipulation the variance of $\hat{\Gamma}$ is resolved into a final general form:

$$\begin{aligned} \text{var}(\hat{\Gamma}) &\approx \left\{ \sum_{i=1}^n \sum_{j=1}^n \left[\left(\frac{\partial \Gamma}{\partial p_{ij}} \right) \Big|_{p_{ij}=\hat{p}_{ij}} \right]^2 \frac{\hat{p}_{ij}}{T} - \sum_{i=1}^n \sum_{j=1}^n \left[\left(\frac{\partial \Gamma}{\partial p_{ij}} \right) \Big|_{p_{ij}=\hat{p}_{ij}} \right]^2 \frac{\hat{p}_{ij}^2}{T} \right\} \\ &- \left\{ \sum_{i=1}^n \sum_{j=1}^n \left[\left(\frac{\partial \Gamma}{\partial p_{ij}} \right) \Big|_{p_{ij}=\hat{p}_{ij}} \right] \sum_{k=1}^n \sum_{l=1}^n \left[\left(\frac{\partial \Gamma}{\partial p_{kl}} \right) \Big|_{p_{kl}=\hat{p}_{kl}} \right] \frac{\hat{p}_{ij}\hat{p}_{kl}}{T} - \sum_{i=1}^n \sum_{j=1}^n \left[\left(\frac{\partial \Gamma}{\partial p_{ij}} \right) \Big|_{p_{ij}=\hat{p}_{ij}} \right]^2 \frac{\hat{p}_{ij}^2}{T} \right\}, \\ \text{var}(\hat{\Gamma}) &\approx \frac{1}{T} \left\{ \sum_{i=1}^n \sum_{j=1}^n \left[\left(\frac{\partial \Gamma}{\partial p_{ij}} \right) \Big|_{p_{ij}=\hat{p}_{ij}} \right]^2 \hat{p}_{ij} - \sum_{i=1}^n \sum_{j=1}^n \left[\left(\frac{\partial \Gamma}{\partial p_{ij}} \right) \Big|_{p_{ij}=\hat{p}_{ij}} \right] \hat{p}_{ij} \sum_{k=1}^n \sum_{l=1}^n \left[\left(\frac{\partial \Gamma}{\partial p_{kl}} \right) \Big|_{p_{kl}=\hat{p}_{kl}} \right] \hat{p}_{kl} \right\} \end{aligned} \quad (38)$$

The first and second derivatives of Γ_1 at the point (p_{11}, \dots, p_{nn}) are as follows:

$$\frac{\partial \Gamma_1}{\partial p_y} = \frac{p_y^2 [1 + \ln(p_y)] - p_y [1 + 2 \ln(p_y)]}{(p_y - 1)^2}, \quad \frac{\partial^2 \Gamma_1}{\partial p_y^2} = \frac{p_y^2 - 4p_y + 2 \ln(p_y) + 3}{(p_y - 1)^3} \quad (39)$$

Substituting the first derivative value for the partial derivative notation in Equation (38) and simplifying gives:

$$\begin{aligned} \text{var}(\hat{\Gamma}_1) \approx \frac{1}{T} \left\{ \sum_{i=1}^n \sum_{j=1}^n \hat{p}_y \frac{[\hat{p}_y^2 (1 + \ln(\hat{p}_y)) - \hat{p}_y (1 + 2 \ln(\hat{p}_y))]^2}{(\hat{p}_y - 1)^4} \right. \\ \left. - \left[\sum_{i=1}^n \sum_{j=1}^n \frac{\hat{p}_y^3 (1 + \ln(\hat{p}_y)) - \hat{p}_y^2 (1 + 2 \ln(\hat{p}_y))}{(\hat{p}_y - 1)^2} \right]^2 \right\} \quad (40) \end{aligned}$$

The first and second derivatives of Γ_2 at the point (p_{11}, \dots, p_{nn}) are as follows:

$$\frac{\partial \Gamma_2}{\partial p_y} = [1 + \ln(p_y)]; \quad \frac{\partial^2 \Gamma_2}{\partial p_y^2} = p_y^{-1}; \quad i, j = 1, 2, \dots, n \quad (41)$$

As before, substituting the first derivative value in Equation (38) and simplifying gives:

$$\text{var}(\hat{\Gamma}_2) \approx \frac{\sum_{i=1}^n \sum_{j=1}^n \hat{p}_y \ln^2(p_y) - [\sum_{i=1}^n \sum_{j=1}^n \hat{p}_y \ln(p_y)]^2}{T} \quad (42)$$

Since RC_2 is just a scaled version of Γ_2 , the variance of \hat{RC}_2 , or $\hat{\Gamma}_{2(\text{scaled})}$, is simply the variance of $\hat{\Gamma}_2$ times the square of the scaling factor, i.e.:

$$\text{var}(\hat{RC}_2) = \text{var}(\hat{\Gamma}_{2(\text{scaled})}) = \frac{\text{var}(\hat{\Gamma}_2)}{4 \ln^2(n)} \quad (43)$$

Consistency

The property of consistency ensures that an estimate is close to the true parameter value with a high probability if the sample size is sufficiently large. Let $\hat{\theta}_T$ be an estimator of θ based on a sample of size T . Sufficient conditions for an estimator $\hat{\theta}_T$ to be consistent for θ are (Judge et al. 1988, p. 85):

$$\lim_{T \rightarrow \infty} E[\hat{\theta}_T] = \theta \quad (44a)$$

$$\lim_{T \rightarrow \infty} \text{var}(\hat{\theta}_T) = 0 \quad (44b)$$

An estimator that satisfies (44a) is said to asymptotically unbiased. Thus an estimator is consistent if any bias it has goes to zero as the sample size grows and if its variance goes to zero as $T \rightarrow \infty$. The limit of Equation (30) is Γ so condition (44a) is met; and the limit of Equations (40), (42), and (43) is zero so condition (44b) is met. Thus $\hat{\Gamma}_1$, $\hat{\Gamma}_2$, and $\hat{\Gamma}_{2(\text{scaled})}$ are all consistent estimators.

Asymptotic Distribution

Kemphorne and Folks (1971, pp. 112-115 and pp. 120-121) have shown how the probabilities for the multinomial distribution tend with increasing sample size to the ordinates for the multivariate normal distribution. They used a moment generating function argument to show that the limiting distribution of the multinomial is the multivariate normal distribution. Because $\hat{\Gamma}_1$, $\hat{\Gamma}_2$, and $\hat{\Gamma}_{2(\text{scaled})}$ are functions of multinomial probabilities, the k -variate normal condition of Equation (12) is met, thus proving asymptotic normality for $\hat{\Gamma}_1$, $\hat{\Gamma}_2$, and $\hat{\Gamma}_{2(\text{scaled})}$.

As an alternative approach to prove that as $T \rightarrow \infty$ the distributions of $\hat{\Gamma}_1$, $\hat{\Gamma}_2$, and $\hat{\Gamma}_{2(\text{scaled})}$ converge to the normal distribution with mean values as in Equation (30) and variances in Equations (40), (42), and (43), write the Taylor expansion shown in Equation (28) in the forms

$$\sqrt{T}(\hat{\Gamma}_1 - \Gamma_1) = \sqrt{T} \left[\sum \sum \frac{(\hat{p}_{ij} - p_{ij}) [p_{ij}^2 [1 + \ln(p_{ij})] - p_{ij} [1 + 2 \ln(p_{ij})]]}{(p_{ij} - 1)^2} + \frac{1}{2} \sum \sum \frac{(\hat{p}_{ij} - p_{ij})^2 [p_{ij}^2 - 4p_{ij} + 2 \ln(p_{ij}) - 3]}{[p_{ij} - \theta(\hat{p}_{ij} - p_{ij}) - 1]^3} \right] \quad (45a)$$

$$\sqrt{T}(\hat{\Gamma}_2 - \Gamma_2) = \sqrt{T} \left[\sum \sum (\hat{p}_{ij} - p_{ij}) [1 + \ln(p_{ij})] + \frac{1}{2} \sum \sum \frac{(\hat{p}_{ij} - p_{ij})^2}{p_{ij}(1 - \theta) - \theta \hat{p}_{ij}} \right] \quad (45b)$$

$$\sqrt{T}(\hat{\Gamma}_{2(\text{scaled})} - \Gamma_{2(\text{scaled})}) = \frac{\sqrt{T}}{2 \ln(n)} \left[\sum \sum (\hat{p}_{ij} - p_{ij}) [1 + \ln(p_{ij})] + \frac{1}{2} \sum \sum \frac{(\hat{p}_{ij} - p_{ij})^2}{p_{ij}(1 - \theta) - \theta \hat{p}_{ij}} \right] \quad (45c)$$

where $0 < \theta < 1$, and apply the convergence in distribution theorem (see definition 4.1 in White 1984, also see theorem 28.4 in Cramér 1946) to the random variables $\sqrt{T}(\hat{\Gamma}_1 - \Gamma_1)$, $\sqrt{T}(\hat{\Gamma}_2 - \Gamma_2)$, and $\sqrt{T}(\hat{\Gamma}_{2(\text{scaled})} - \Gamma_{2(\text{scaled})})$. The theorem as stated in Cramér says that if you have a function $H(m_v, m_q)$ of two central moments, then “If, in some neighbourhood of the point $m_v = \mu_v$, $m_q = \mu_q$, the function $H(m_v, m_q)$ is continuous and has continuous derivatives of the first and second order with respect to the arguments m_v and m_q , the random variable $H(m_v, m_q)$ is asymptotically normal” It follows from this theorem that any sample characteristic based on moments is, for large values of T , approximately normally distributed about the corresponding population characteristic.

I have just computed the mean value and variance for three statistical estimates, $\hat{\Gamma}_1$, $\hat{\Gamma}_2$, and $\hat{\Gamma}_{2(\text{scaled})}$, of the contagion of a landscape. The estimates were shown to be asymptotically unbiased, consistent, and asymptotically normal.

A direct and very important use of the asymptotic normality and variance of the given estimators is in hypothesis testing. Tests against an *a priori* null can be accomplished via a one-sample *t*-test or construction of a confidence interval about the point estimate. Likewise, two point estimates can be compared using a two-sample *t*-test. Three or more estimates can be compared using analysis of variance. Often, hypotheses of interest can be expressed in terms of linear combinations of the estimates as $\mathbf{R}\hat{\boldsymbol{\Gamma}} = \mathbf{r}$ where \mathbf{R} and \mathbf{r} are a matrix and a vector of known elements that specify the hypotheses of interest. For example, if the hypothesis is that the elements of $\hat{\boldsymbol{\Gamma}}$ sum to unity, $\mathbf{R} = [1, 1, \dots, 1]$ and $\mathbf{r} = 1$. The Wald, Lagrange multiplier, and quasi-likelihood ratio tests can be used to test the null hypothesis $\mathbf{R}\hat{\boldsymbol{\Gamma}} = \mathbf{r}$ versus the alternative $\mathbf{R}\hat{\boldsymbol{\Gamma}} \neq \mathbf{r}$.

MATERIALS AND METHODS

Behavior of Γ_1 , Γ_2 , and RC_2

I followed two approaches to formally investigate the behavior of the new Γ_1 and Γ_2 indices, and RC_2 . The first approach was similar to one used by Li and Reynolds (1993).

Increasing Gradient of Evenness

I generated a series of simulated landscapes with different spatial configurations (random, uniform, and clustered) and numbers of patch types (from 2 to 10), with an increasing gradient of evenness of the proportions of patch types. A simple 0 to 1 scaled measure of evenness is given by (Turner 1989)

$$RE = \frac{-\ln\left(\sum_{i=1}^n p_i^2\right)}{\ln(n)} \quad (46)$$

where p_i is the proportion of the landscape in patch type i and RE (relative evenness) approaching 0 means increasing unevenness of the n categories and $RE = 1$ means all categories occur in equal proportion. The simulated spatial maps were generated with software (program SHAPC) written and provided by Dr. Harbin Li¹ of Duke University. For details on the algorithms used in program SHAPC refer to Li et al. (1993) and Li and Reynolds (1994). In particular, a random landscape is generated by filling pixels left to right then top to bottom using a random number generator. A uniform landscape is generated by

¹Present address: USDA Forest Service, Southern Research Station, 2730 Savannah Highway, Charleston, SC 29414.

uniform spacing of square blocks on the map. An aggregated landscape is generated by randomly placing regularly shaped clusters on the map.

For each of the three configurations, nine maps were generated. The first map has two patch types, simply denoted 1 and 2, covering 90% and 10%, respectively. The second map has three patch types, 1, 2, and 3, covering 80%, 10%, and 10%, respectively. The third map has patch type 1 covering 70% and the remaining three types covering 10% each, and so on. The ninth map contains 10 patch types each covering 10% of the surface area. The rationale behind this scheme is to have a coverage of points between the extremes of high contagion to low contagion. Conceptually, high contagion is the case where one patch type (or a small number) dominates the landscape, occupying most of the area. Conversely, low contagion occurs when a landscape has many patch types occupying roughly the same amount of area. Spatial pattern also affects contagion. Conceptually, a random pattern (a fragmented landscape) means less contagion and an aggregated pattern results in more contagion.

Each map was 20×20 squares for a total of 400 squares. Index values were computed based on counting adjacencies (rook's rule) around all 400 squares on each map. I wrote a FORTRAN program, CONTAGION, which reads a categorical map input as a matrix of integers (each integer value representing a different category) and computes Γ_1 , Γ_2 , RC_2 , their variances, and RE . The source code along with an example input file and program output are listed in Appendix B.

The outlined scheme resulted in 27 simulated landscapes, nine for each spatial configuration. For each index, the computed values (calculated with program

CONTAGION) were plotted over number of patch types by spatial configuration, which gives a graph showing three lines, each determined by nine points. The plots show the sensitivity of the new indices to changing numbers of categories/evenness (i.e., composition) and spatial configurations. The index RC_2 was computed on the same 27 maps and values plotted for comparison with the new indices.

Same Degree of Evenness

As a second approach to looking at the behavior of the new indices, I generated another series of simulated landscapes for random, uniform, and clustered spatial configurations and numbers of patch types (from 2 to 10), but with completely even proportions of patch types. That is, for two patch types each type occupied $\frac{1}{2}$ of the area; for three patch types each type occupied $\frac{1}{3}$ of the area, and so on ($RE \approx 1$ for each map). Again, this resulted in 27 simulated landscapes, nine for each spatial configuration. A plot of the index values computed on these maps show if the new indices, proportions being equal, are sensitive to increasing number of categories under different spatial configurations. The index RC_2 was also computed on these same 27 maps and values plotted for comparison with the new indices.

Monte Carlo Investigation of Distribution

Using program SHAPC, I generated three simulated $24 \times 24 = 576$ pixel landscapes; a random, a uniform, and a clustered, each with five patch types. The landscapes were constructed such that three patch types each occupied approximately 15% of the area, one patch type occupied 35% and the last patch type covered 20%. I used program CONTAGION to compute contagion values and their corresponding moment-based

variances for the three landscapes. The method of Sherman and Carlstein (1994) was used to compute nonparametric estimates of contagion variance for the three landscapes.

I wrote a FORTRAN program, VAREST, as a means to systematically subdivide a landscape D (input as a matrix of integers) into nonoverlapping subrectangles D_k with corresponding sublattice data Z_{D_k} for computation of the nonparametric variance using Equation (18). The subrectangle size must be specified in the program. Appendix C lists the source code for program VAREST, and gives an example input file and shows program output.

As an aid to determine an appropriate subrectangle size, a sequence of runs was conducted with each of the three simulated landscapes specifying different subrectangle sizes. Uniformity of RE across the K subregions (where $K = \text{matrix size} \div \text{subrectangle size}$) was used as the selection criteria. A subsquare size of 12×12 pixels was selected, thus each landscape was divided into $24^2/12^2 = 4$ congruent subregions for variance estimation. The nonparametric variances were compared against the theoretical moment-based values as a validity check.

Replicate histograms (Sherman and Carlstein 1996) were used as a means to nonparametrically study the sample distribution of contagion. The three simulated landscapes were each divided into K overlapping subsquare replicates. Empirical distributions were constructed via Equation (19) and kernel smoothed histograms were plotted. The histogram plots show the general shape of the sampling distribution of contagion. This provides a validity check against the theoretical asymptotic distribution.

I wrote a FORTRAN program, REPHIST, as a means to systematically subdivide a landscape D (input as an $r \times c$ matrix of integers) into all possible overlapping subsquares D_l^k (where l represents the length of a side, i.e., squares are $l \times l$) with corresponding sublattice data $Z_{D_l^k}$ for construction of an empirical distribution. The subsquare size l must be specified in the program. As with program VAREST, a trial process was used to determine an appropriate subsquare size. Based on uniformity of RE across the K subregions, a subsquare size of 12×12 pixels was selected. Each histogram was therefore constructed from $K = (r - l + 1)(c - l + 1)$ or $(24 - 12 + 1)(24 - 12 + 1) = 169$ values. Appendix D lists the source code for program REPHIST, and gives an example input file and shows program output.

Data

Physiographic Provinces and Forest Cover Type Data

To assess landscape habitat diversity one must necessarily specify the spatial boundaries of a landscape chosen for study. For geographically targeted units, the boundaries may be thought of as analogous to those of an *island*; the word, used in an ecological sense, connotes any environment entirely surrounded by another with strikingly different properties, the two being separated by a fairly abrupt boundary (Pielou 1975). Major patterns of landform-geologic material or “physiographic provinces” are well defined in Alabama (Hodgkins et al. 1979) (Figure 2). This landscape-level scale is useful for examination of spatial patterns of forest cover diversity.

The Southern Forest Inventory and Analysis (SFIA) unit of the USDA Forest Service, Southern Research Station, conducts continuing inventories of forest resources in

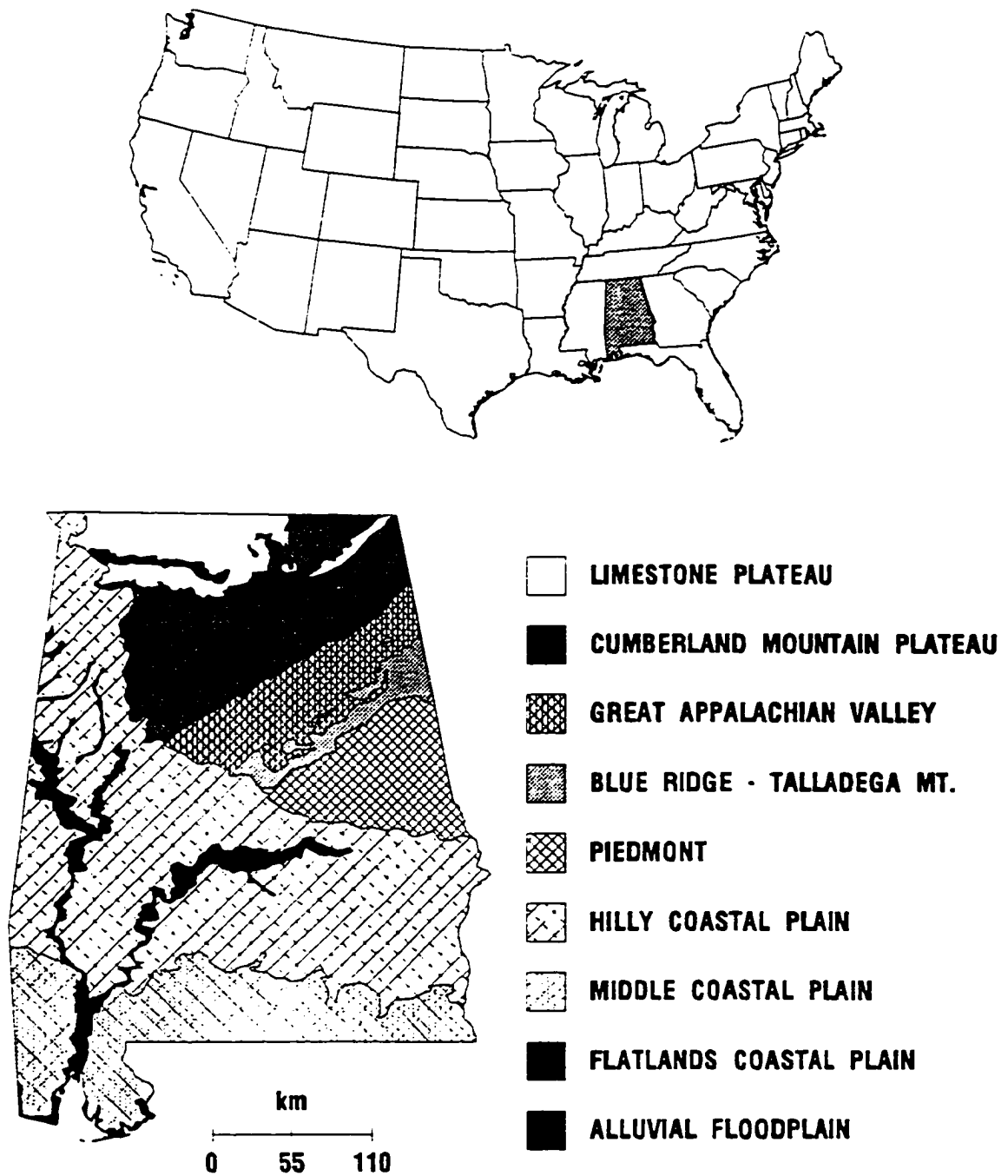


Figure 2. The nine physiographic provinces of Alabama.

thirteen Midsouth States. Data are collected from trees occurring on sample plots spaced across each State on a 3- by 3-mile (4.8- by 4.8-km) grid. From these tree data a forest cover type is identified for each plot. General forest cover types are listed in Table 1. Detailed descriptions of the SFIA data can be found in May (1990). The survey periods 1972, 1982, and 1990 are available in digital format. The SFIA data are easily stratified by physiographic province. This provides three points in time for examining trends in forest cover type diversity on well defined landscape-scale units.

It should be mentioned that as with all data, USDA Forest Service Forest Inventory and Analysis data have pros and cons. On the pro side side, (i) forest surveys date back to the 1930's, (ii) survey data are readily available over the internet, and (iii) they cover the entire U.S. On the con side, (i) all nonforest land is classified the same, (ii) for some uses, especially GIS, resolution is low because plots are spaced at 3-mile (4.8-km) intervals, (iii) as with all sample data, point estimates of variables (e.g., proportions of cover types) have statistical error.

Study Areas

To illustrate use of the new indices on real-world data, three adjacent physiographic provinces in Alabama were chosen for study: (i) Great Appalachian Valley Province, (ii) Blue Ridge-Talladega Mountain Province, and (iii) Piedmont Province (see Figure 2). Locations of survey plots within the three physiographic provinces are shown in Figure 3. The SFIA survey plots are one of many possible data realizations. A forest cover type was determined at each of these plots.

Table 1. USDA Forest Service general forest cover types¹ occurring on three physiographic provinces in the state of Alabama, USA.

General forest cover types	Species
Longleaf-slash pine	<i>Pinus palustris</i> - <i>Pinus elliotii</i>
Loblolly-shortleaf pine	<i>Pinus taeda</i> - <i>Pinus echinata</i>
Oak-pine	<i>Quercus</i> sp.- <i>Pinus</i> sp.
Oak-hickory	<i>Quercus</i> sp.- <i>Carya</i> sp.
Oak-gum-cypress	<i>Quercus</i> sp.- <i>Liquidambar styraciflua</i> - <i>Taxodium distichum</i>
Elm-ash-cottonwood	<i>Ulmus</i> sp.- <i>Fraxinus</i> sp.- <i>Populus</i> sp.
Nontyped	
Nonforest	

¹Source: May (1990, p. 6)

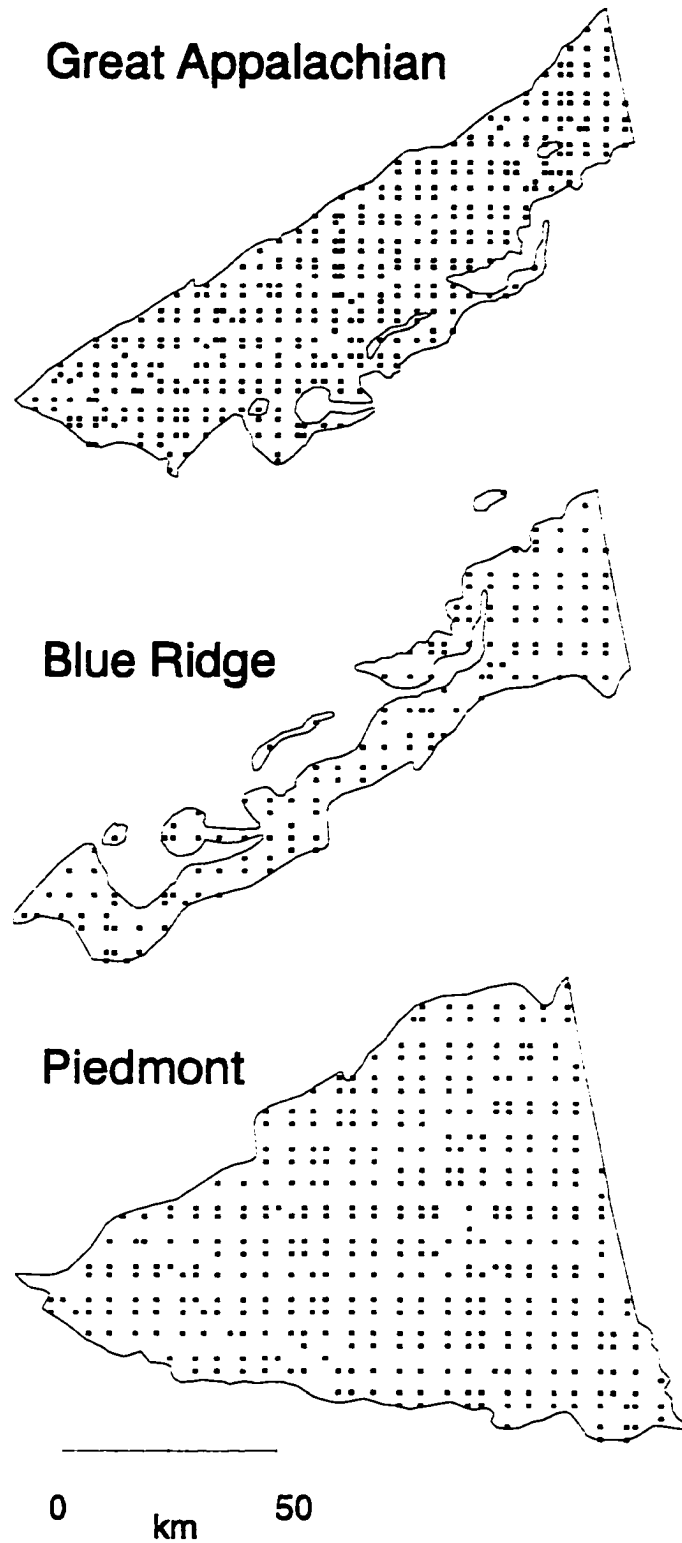


Figure 3. Distribution of Southern Forest Inventory and Analysis survey plots on three physiographic provinces in Alabama.

Thiessen polygons were used to apportion a point coverage into regions known as Thiessen or proximal polygons (Environmental Systems Research Institute 1992). Each region contains only one point and has the unique property that any location within a region is closer to the region's point than to the point of any other region. Each physiographic province point coverage was apportioned into Thiessen polygons and like polygons (polygons of the same forest cover type) were marked the same to create a landscape level view of forest cover types for each province at the three survey years. Again it is important to remember that the Thiessen polygons represent a data realization view. A polygon is superimposed on top of each plot to establish join counts. Figure 4 is a landscape level view of cover types on the Great Appalachian Valley Province; Figure 5 is a landscape level view of cover types on the Blue Ridge-Talladega Mountain Province; and Figure 6 is a landscape level view of cover types on the Piedmont Province. Proportions of each cover type on each province by survey year are given in Table 2.

Hypotheses and Hypothesis Testing

It is generally believed that prior to the 1900's forested landscapes in the South were more homogeneous and contiguous than today. Exploitative logging, agriculture, forest type conversion, and other factors have altered, and continue to alter, the mosaic of forest cover types on the landscape. Today landscape flux (changes in composition and/or configuration) can possibly occur on the time scale of a decade. Therefore, it is natural to hypothesize that landscape flux is occurring on the physiographic provinces. Also, it is of general interest to compare physiographic provinces for similarity or differences in contagion. Using the new indices, contagion values (C) were computed for each province

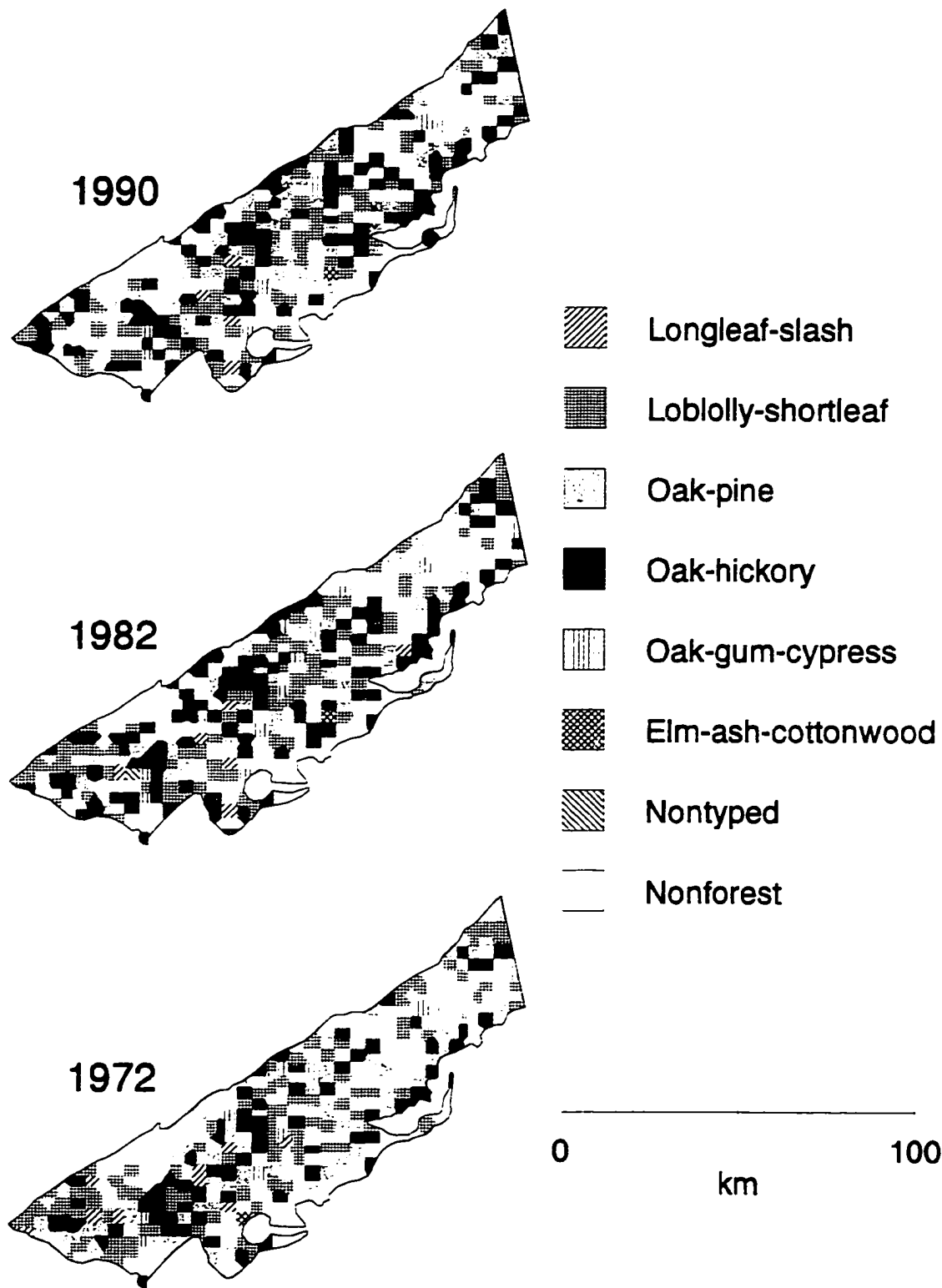


Figure 4. Forest Cover type changes through time in the Great Appalachian Valley physiographic province based on Southern Forest Inventory and Analysis survey data.

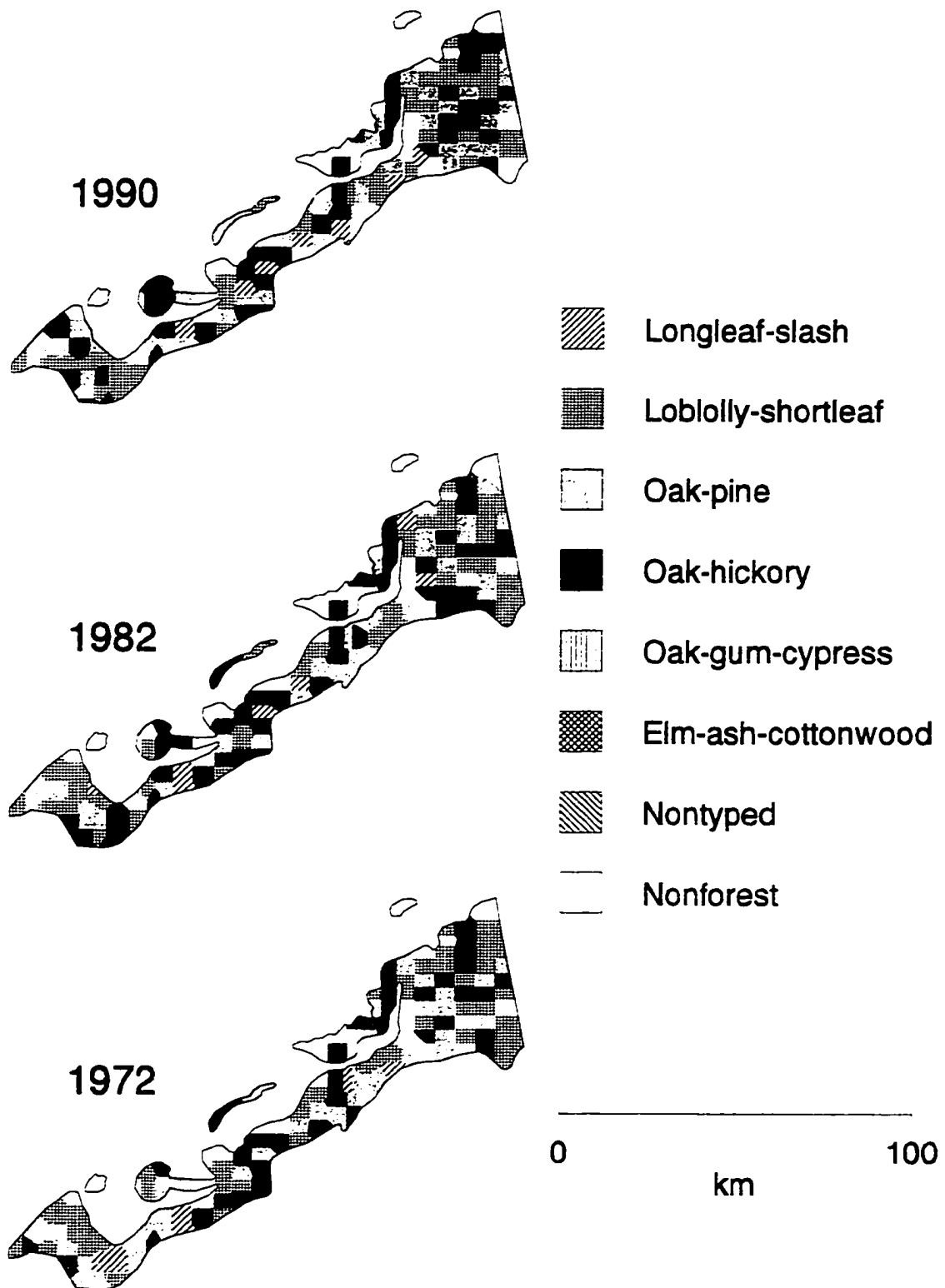


Figure 5. Forest Cover type changes through time in the Blue Ridge-Talladega Mountain physiographic province based on Southern Forest Inventory and Analysis survey data.

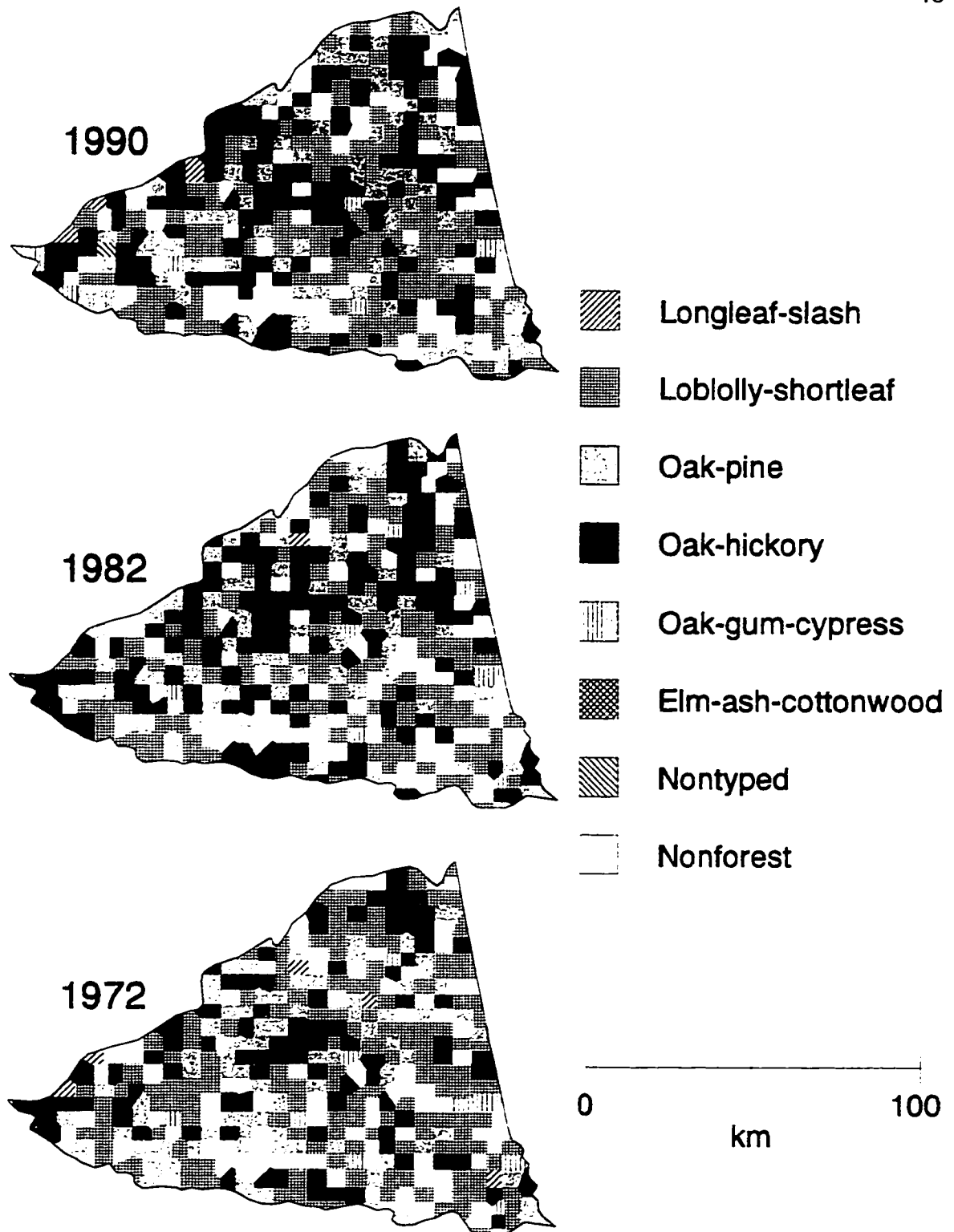


Figure 6. Forest Cover type changes through time in the Piedmont physiographic province based on Southern Forest Inventory and Analysis survey data.

Table 2. Proportion of each forest cover type on each province by survey year.

Forest cover type	Great Appalachian Valley			Blue Ridge - Talladega Mt.			Piedmont		
	72	82	90	72	82	90	72	82	90
----- values in percent -----									
Longleaf-slash pine	2.2	1.5	1.0	3.9	4.5	4.5	1.2	0.2	0.7
Loblolly-shortleaf pine	22.4	18.9	18.2	26.5	24.5	22.6	34.5	26.4	27.4
Oak-pine	19.2	12.5	15.5	23.9	18.7	23.9	20.0	15.6	18.8
Oak-hickory	16.7	26.3	24.8	24.5	31.0	27.7	21.8	33.3	32.0
Oak-gum-cypress	2.9	3.2	2.0		0.6		2.2	1.7	1.7
Elm-ash-cottonwood	0.2	0.2	0.2						
Nontyped		0.2							0.2
Nonforest	36.4	37.1	38.3	21.3	20.6	21.3	20.3	22.7	19.1

for the survey periods 1972, 1982, and 1990. Let A represent the Great Appalachian Valley Province, let B represent the Blue Ridge-Talladega Mountain Province, and let P represent the Piedmont Province. The following constitute a logical set of hypotheses for the landscape data:

$$1) \text{ For Appalachian, } H_0: C_{72}=C_{82}=C_{90}$$

$$2) \text{ For Blue Ridge, } H_0: C_{72}=C_{82}=C_{90}$$

$$3) \text{ For Piedmont, } H_0: C_{72}=C_{82}=C_{90}$$

$$4) \text{ For 1972, } H_0: C_A = C_B = C_P$$

$$5) \text{ For 1982, } H_0: C_A = C_B = C_P$$

$$6) \text{ For 1990, } H_0: C_A = C_B = C_P$$

Variance formulas have been derived for the two new indices, and the distribution of both indices is asymptotically normal, hence hypothesis testing can be accomplished through application of one-way analysis of variance (ANOVA). One-way ANOVA's can be constructed as follows: let t = number of groups, and T_i = number of observations for the i th group, then

$$\text{among groups variance} = \sum_{i=1}^t (C_i - \bar{C})^2 / (t - 1); \quad \bar{C} = \sum_{i=1}^t T_i C_i / \sum_{i=1}^t T_i \quad (47a)$$

$$\text{within groups variance} = \sum_{i=1}^t T_i \text{var}(C_i) / \sum_{i=1}^t T_i \quad (47b)$$

$$F = \frac{\text{among groups variance}}{\text{within groups variance}} \quad \text{with } t - 1, \sum T_i - t \text{ degrees of freedom} \quad (47c)$$

The formula just given for within groups variance is a pooled estimate of the common variance. If variances are not equal then weighted ANOVA should be used. In such a case,

divide each index value by the square root of its variance and compute among groups variance. Within groups variance is now 1, hence F reduces to among groups variance on the weighted index values.

The degrees of freedom specified in (47c) assumes the F statistic is based on a ratio of independent χ^2 values divided by their degrees of freedom. In particular the denominator is based on asymptotic variances. White (1984) in Theorem 4.30 (p. 71) shows that if one has a consistent estimator of the variance \hat{v}_T such that $\hat{v}_T - v_T \xrightarrow{p} 0$ then an asymptotic χ^2 distribution results. As an alternative one could use infinite degrees of freedom for the denominator of F in (47c).

The tests of hypotheses were conducted using $\alpha=.10$. This level was chosen to increase power of the test. The results of the hypothesis tests were used to draw general conclusions about forest cover type diversity on the three physiographic provinces and use of the new indices.

RESULTS AND DISCUSSION

Behavior of Γ_1 , Γ_2 , and RC_2

Increasing Gradient of Evenness

Figure 7 contains the nine simulated random landscapes for 2 to 10 patch types along an increasing gradient of evenness. Figure 7 shows there is a problem with the random landscape generator in program SHAPC, as only part of the space is being randomly filled, proceeding from left to right. Figures 8 and 9 show the simulated uniform and aggregated landscape maps for 2 to 10 patch types with a concomitant increase in patch evenness. Index values for Γ_1 , Γ_2 , and RC_2 were computed on these 27 maps, along with RE (relative evenness of the patch types), and are listed in Appendix E, Table E1.

A graph of the Γ_1 values from Table E1 is displayed as Figure 10. As is readily seen in Figure 10, Γ_1 is sensitive to both composition and configuration. The three spatial configurations separate out logically with aggregated landscapes having the highest values, followed by uniform landscapes, and then the randomly arranged landscapes having the lowest values. This is as expected since random landscapes have little spatial autocorrelation whereas uniform and aggregated landscapes have increasing spatial autocorrelation of patches. There is a sharp decrease in all three curves with increasing number of patches and evenness, covering nearly the full range of index values. This meets with the conceptual definition of contagion.

A graph of the Γ_2 values from Table E1 is displayed as Figure 11. The index Γ_2 shows sensitivity to composition and configuration for these data. The three spatial

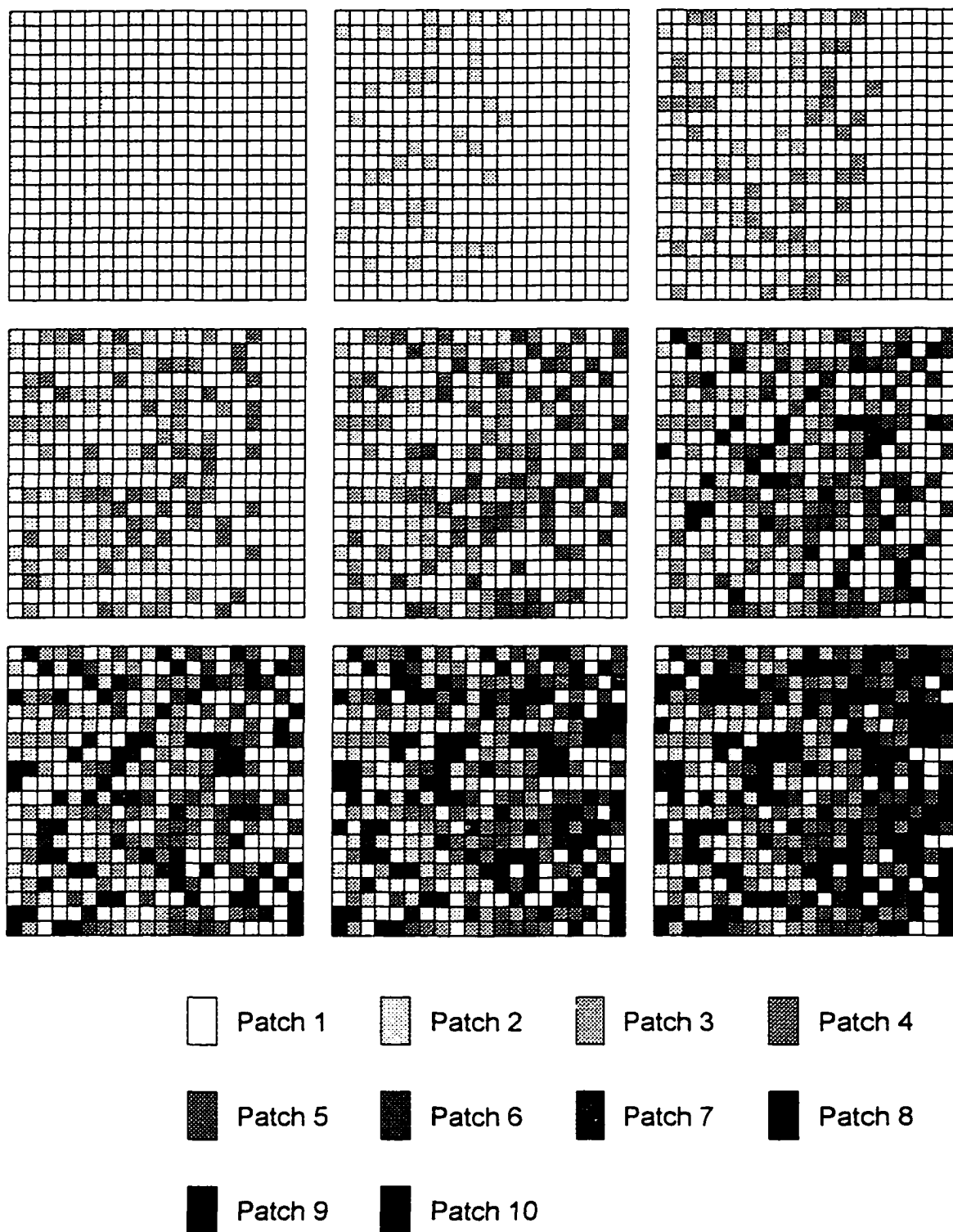


Figure 7. Simulated random landscape maps for 2 to 10 patch types with the base patch type (white) coverage decreasing by 10% with each new patch type. The first map has a coverage of 90% for the base patch type and 10% for the second patch type.

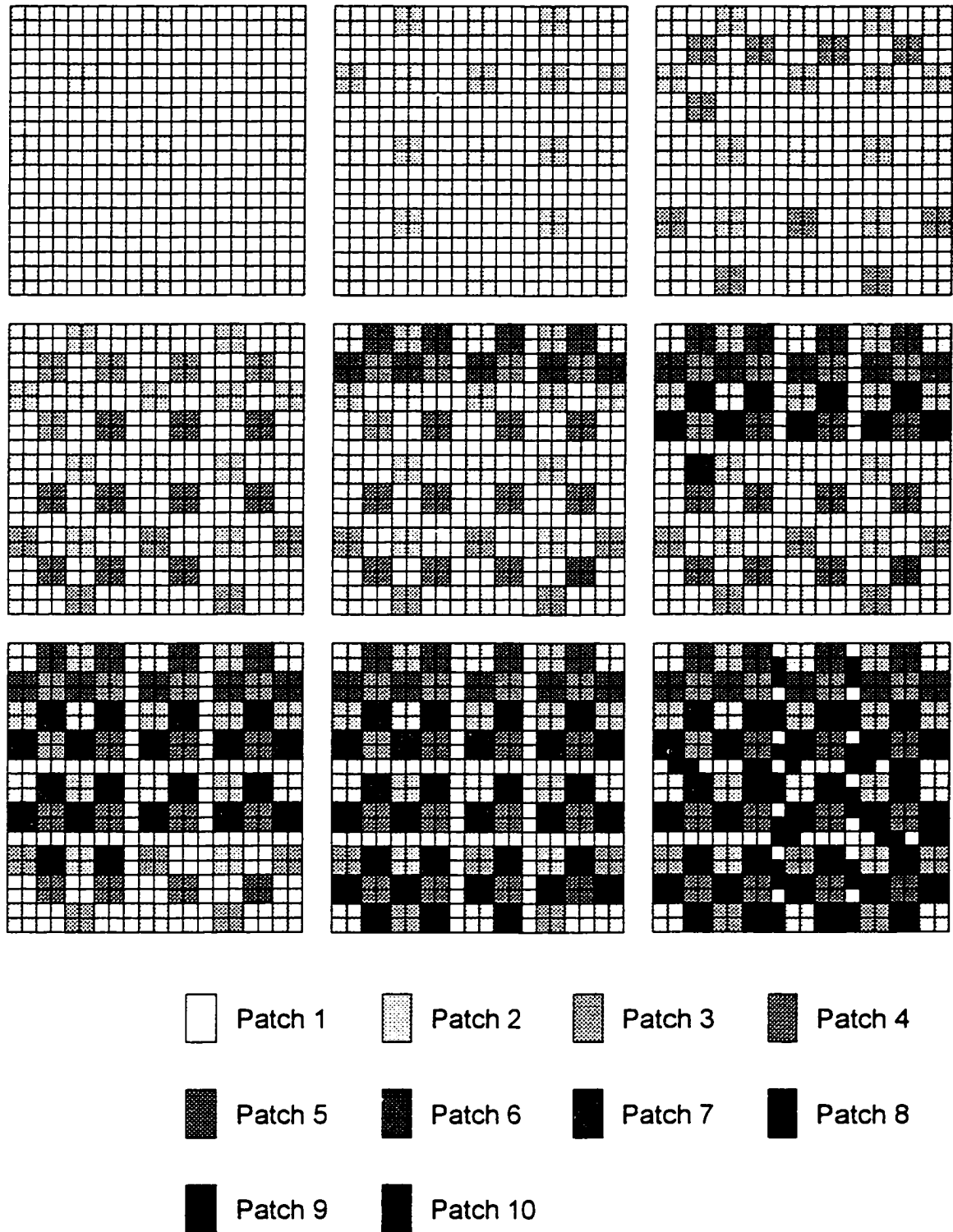


Figure 8. Simulated uniform landscape maps for 2 to 10 patch types with the base patch type (white) coverage decreasing by 10% with each new patch type. The first map has a coverage of 90% for the base patch type and 10% for the second patch type.

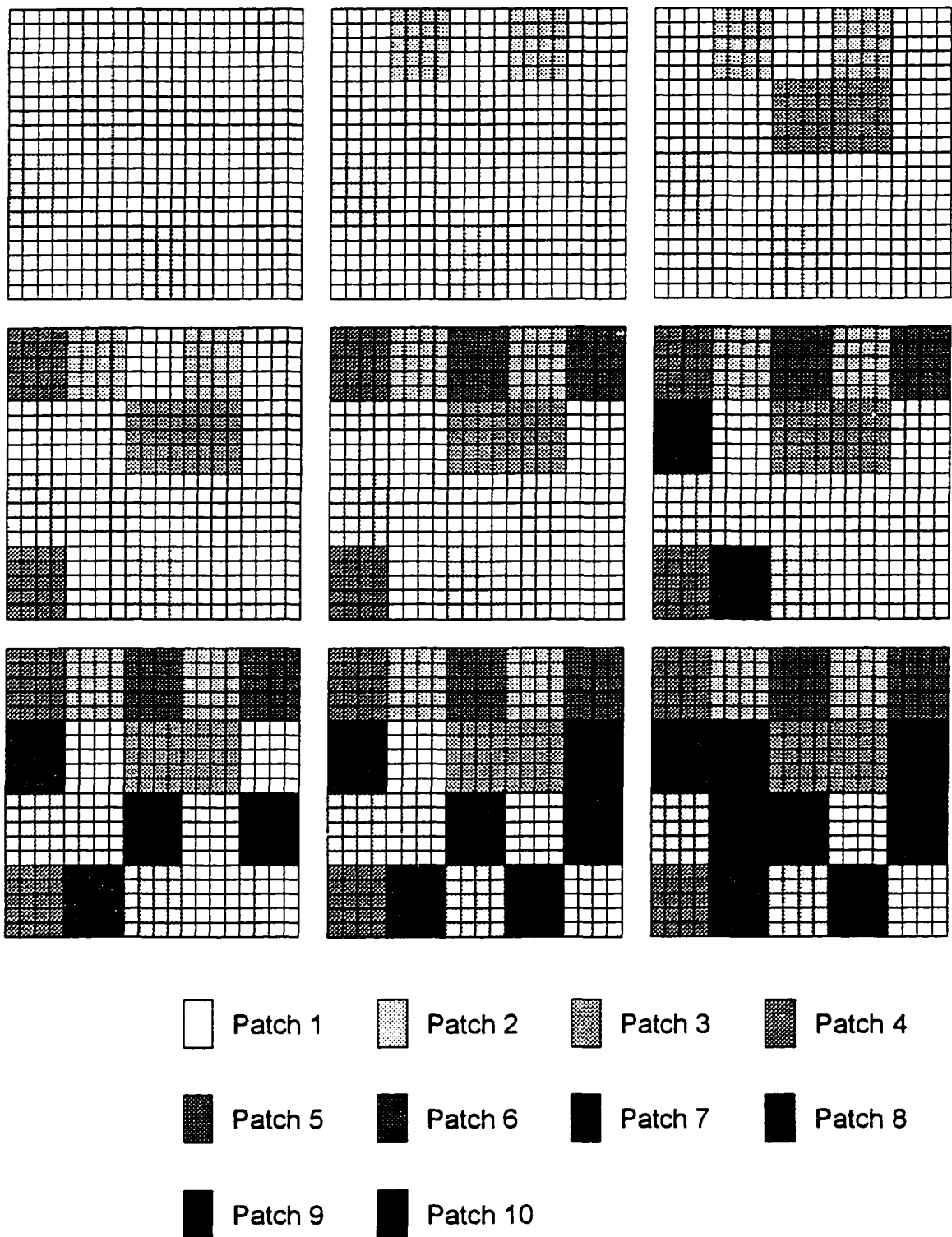


Figure 9. Simulated aggregated landscape maps for 2 to 10 patch types with the base patch type (white) coverage decreasing by 10% with each new patch type. The first map has a coverage of 90% for the base patch type and 10% for the second patch type.

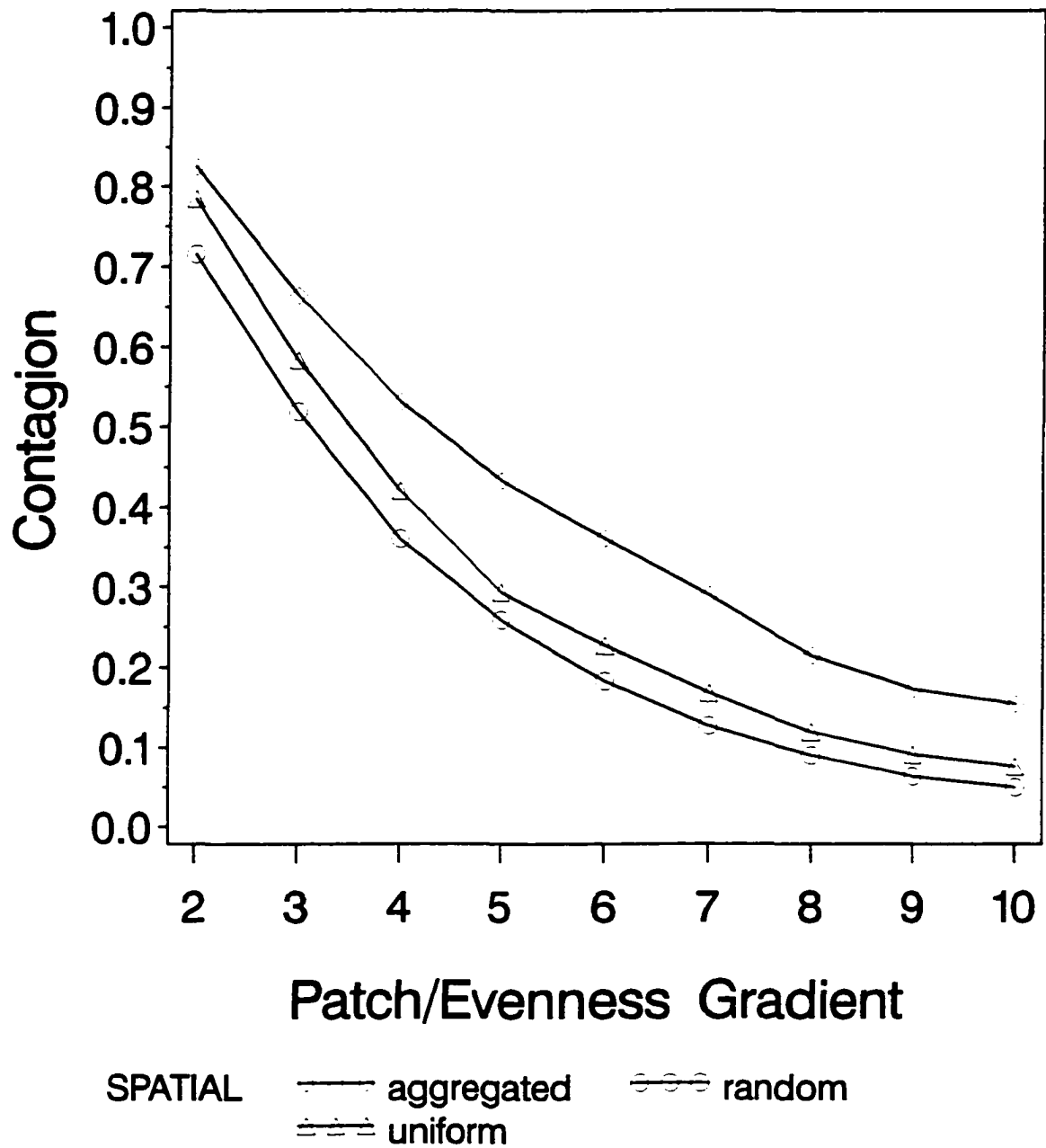


Figure 10. Trend lines showing relationship between Γ_1 and the two controlled variables: spatial pattern and number of patch types. There is an increasing gradient of evenness along the x-axis.

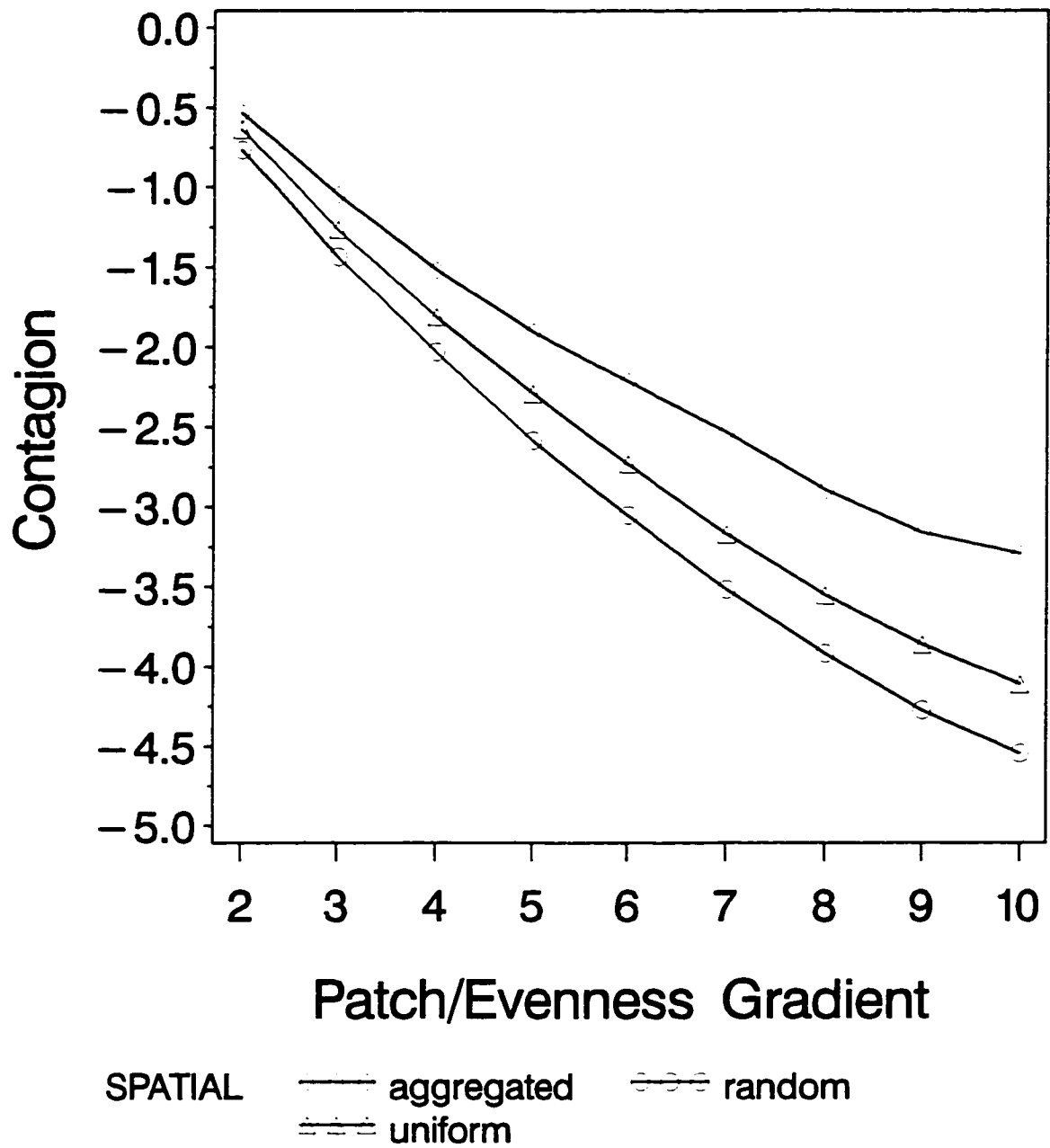


Figure 11. Trend lines showing relationship between Γ_2 and the two controlled variables: spatial pattern and number of patch types. There is an increasing gradient of evenness along the x-axis.

configurations separate out logically with aggregated landscapes having the highest values, followed by uniform landscapes, and then the randomly arranged landscapes having the lowest values. The lines have less curvature than in Figure 10 giving slightly better separation between index values at the higher patch/evenness gradient values.

Figure 12 displays a graph of the RC_2 values. The three spatial configurations separate out very cleanly and logically, but the values only fall between 0 and 0.6. Thus the RC_2 values do not cover most of the 0 to 1 range, in contrast to Γ_1 . The RC_2 curves do not slope as much as the Γ_1 curves or the Γ_2 curves, so RC_2 does not appear to be as sensitive to composition as either Γ_1 or Γ_2 for these landscapes.

Same Degree of Evenness

Figure 13 contains the nine simulated random landscapes for 2 to 10 patch types with $RE \approx 1$ for each map. Likewise Figures 14 and 15 show the simulated uniform and aggregated landscape maps for 2 to 10 patch types with $RE \approx 1$ for each map. Index values for Γ_1 , Γ_2 , and RC_2 , computed on these 27 maps, are listed in Appendix E, Table E2.

A graph of the Γ_1 values from Table E2 is displayed as Figure 16. Again we see that Γ_1 distinguishes between the three spatial configurations and decreases with increases in number of land cover categories. Its behavior is consistent with changes in composition and configuration.

Figure 17 displays a graph of the Γ_2 values. Like Γ_1 , Γ_2 distinguishes between the three spatial configurations and decreases with increases in number of land cover categories. The lines in Figure 17 have less curvature than in Figure 16, hence they flatten out less giving slightly better separation between index values at the higher number of patch types.

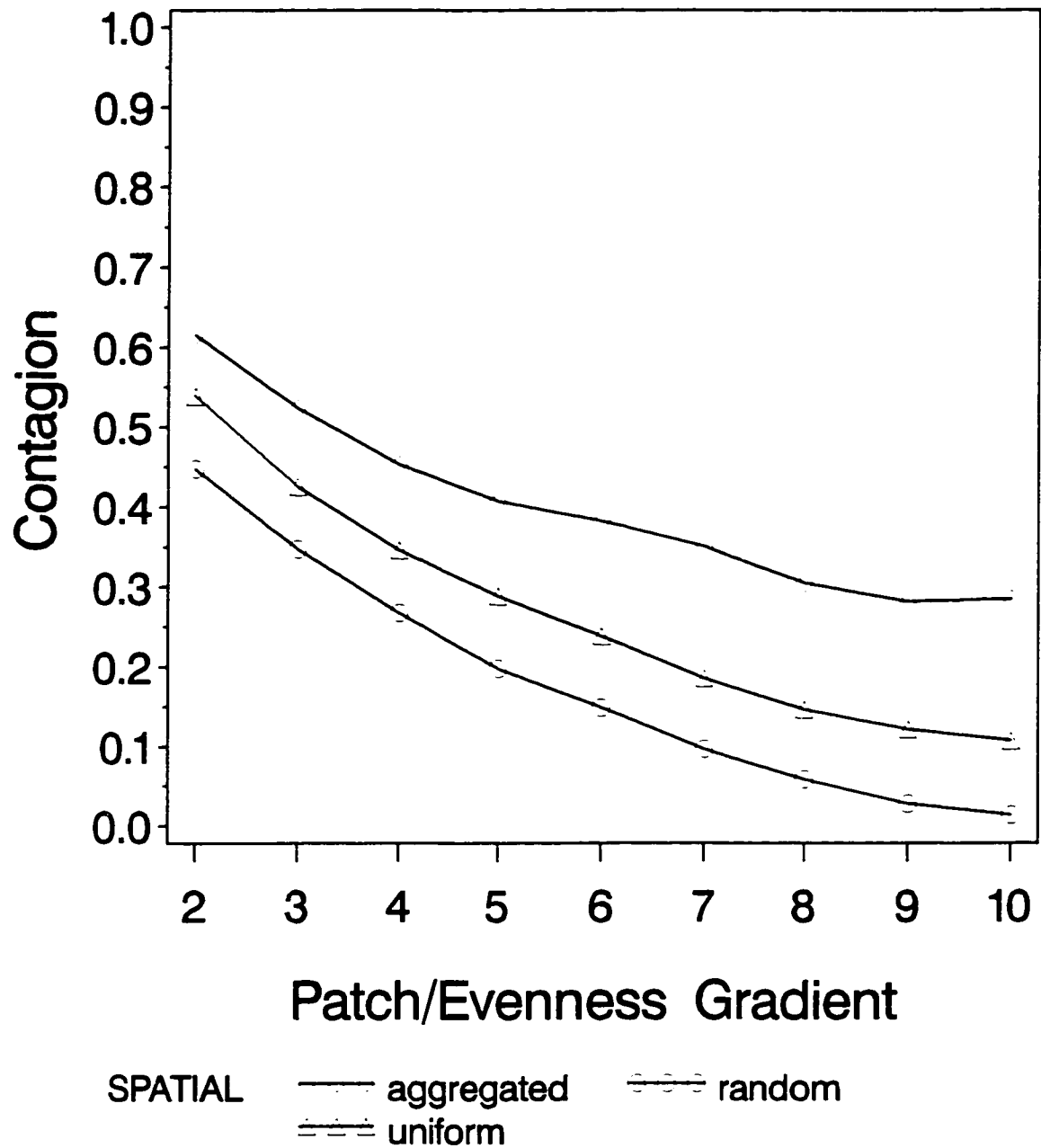


Figure 12. Trend lines showing relationship between RC_2 and the two controlled variables: spatial pattern and number of patch types. There is an increasing gradient of evenness along the x-axis.

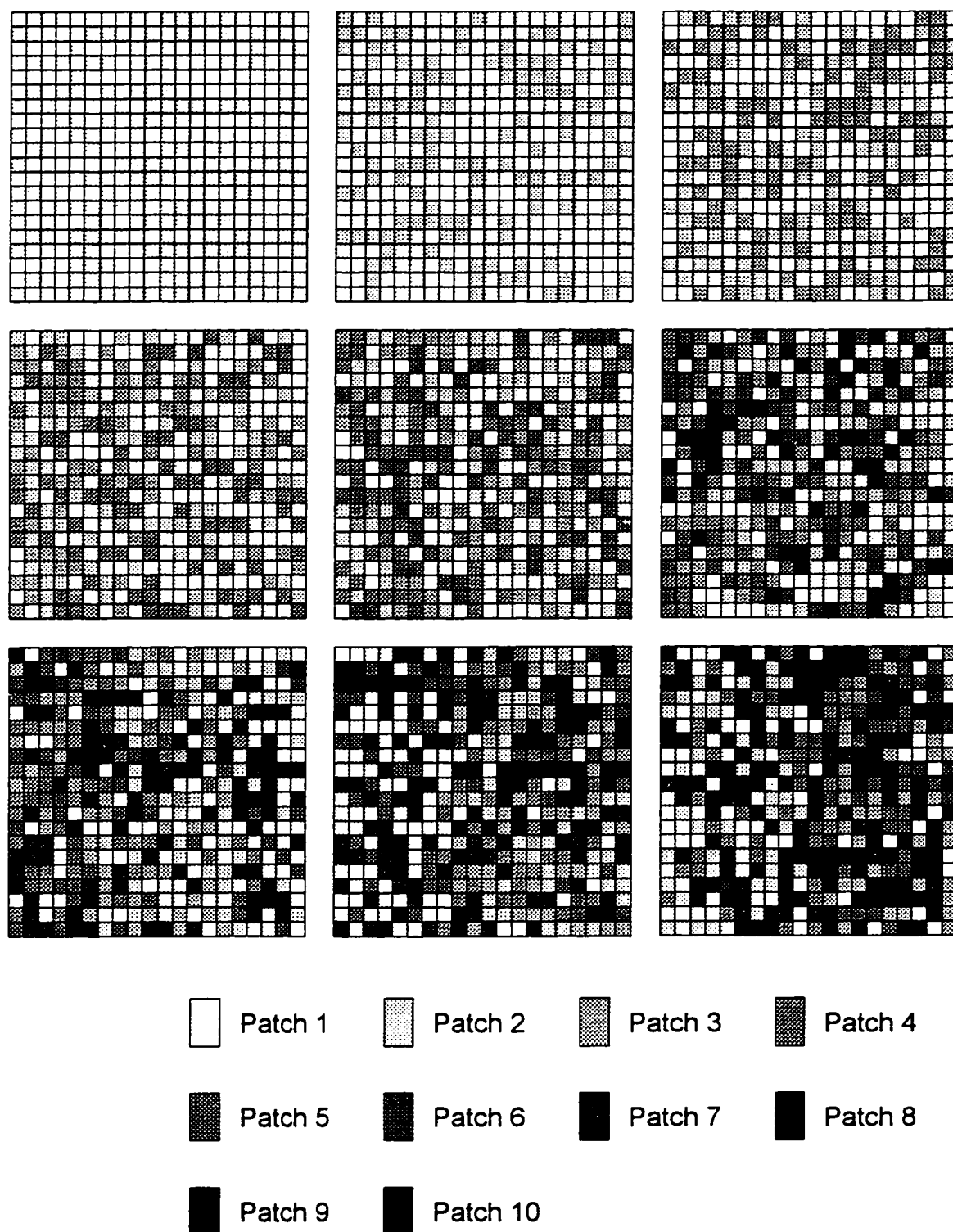


Figure 13. Simulated random landscape maps for 2 to 10 patch types with all patch types on each map having equal coverage.

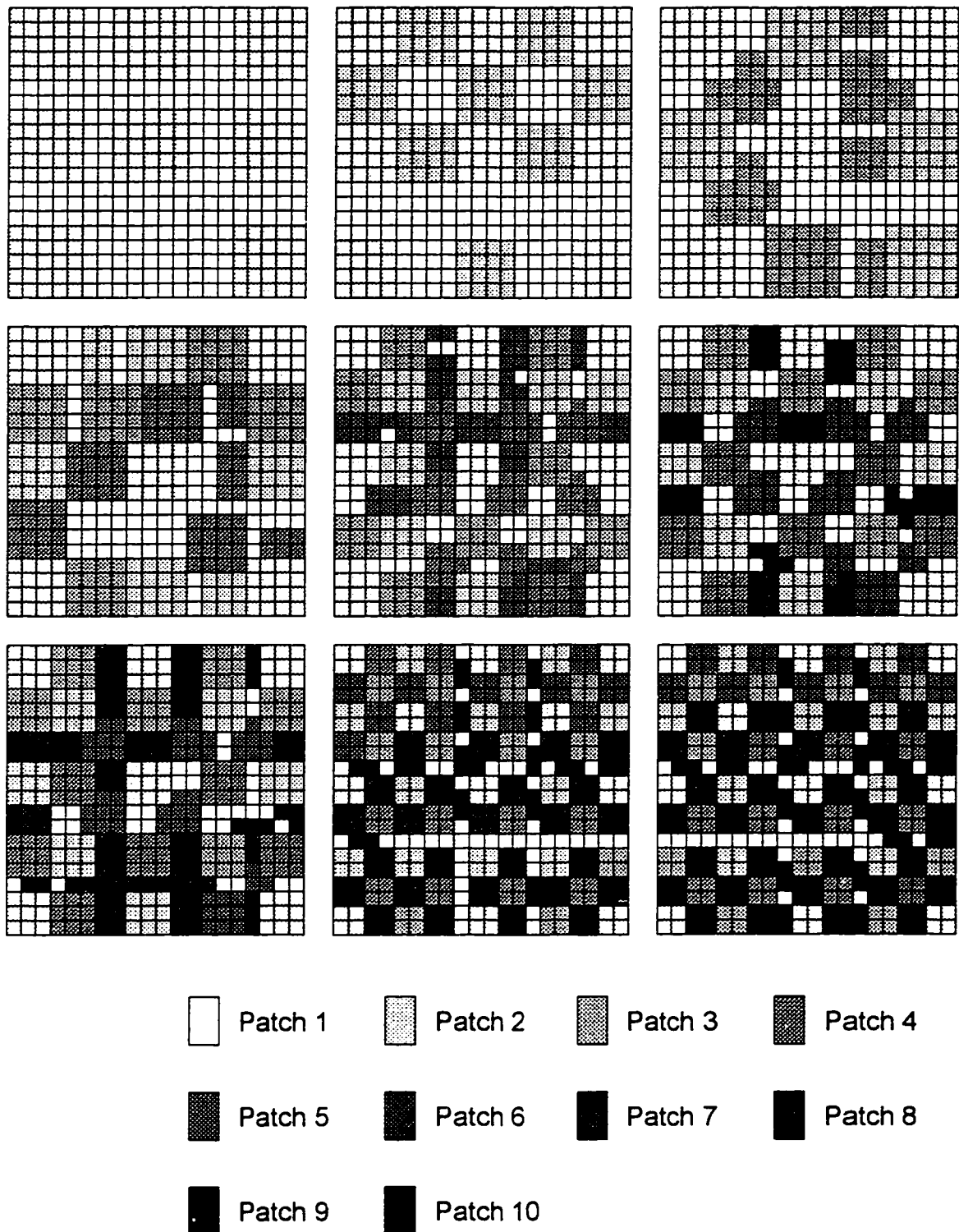


Figure 14. Simulated uniform landscape maps for 2 to 10 patch types with all patch types on each map having equal coverage.

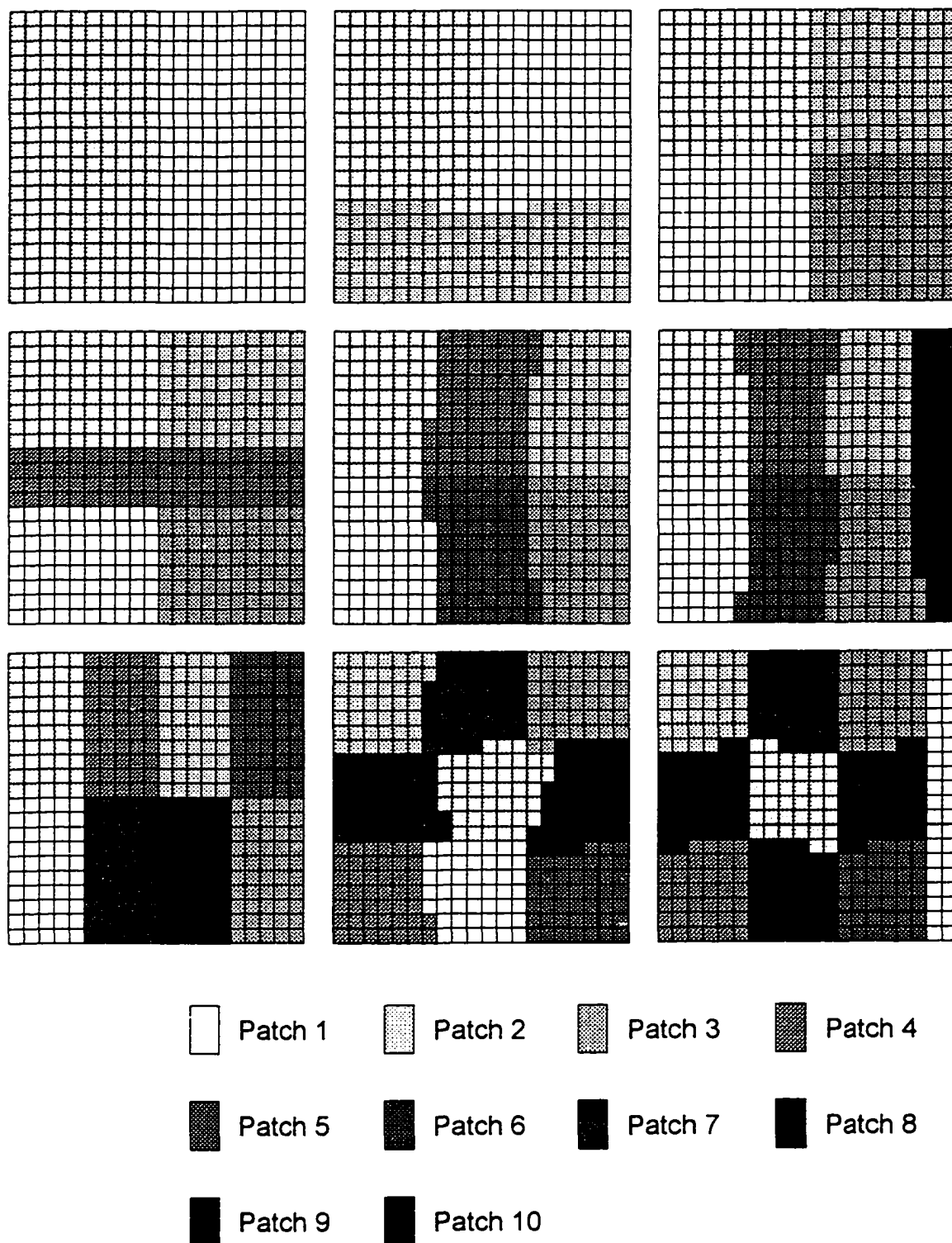


Figure 15. Simulated aggregated landscape maps for 2 to 10 patch types with all patch types on each map having equal coverage.

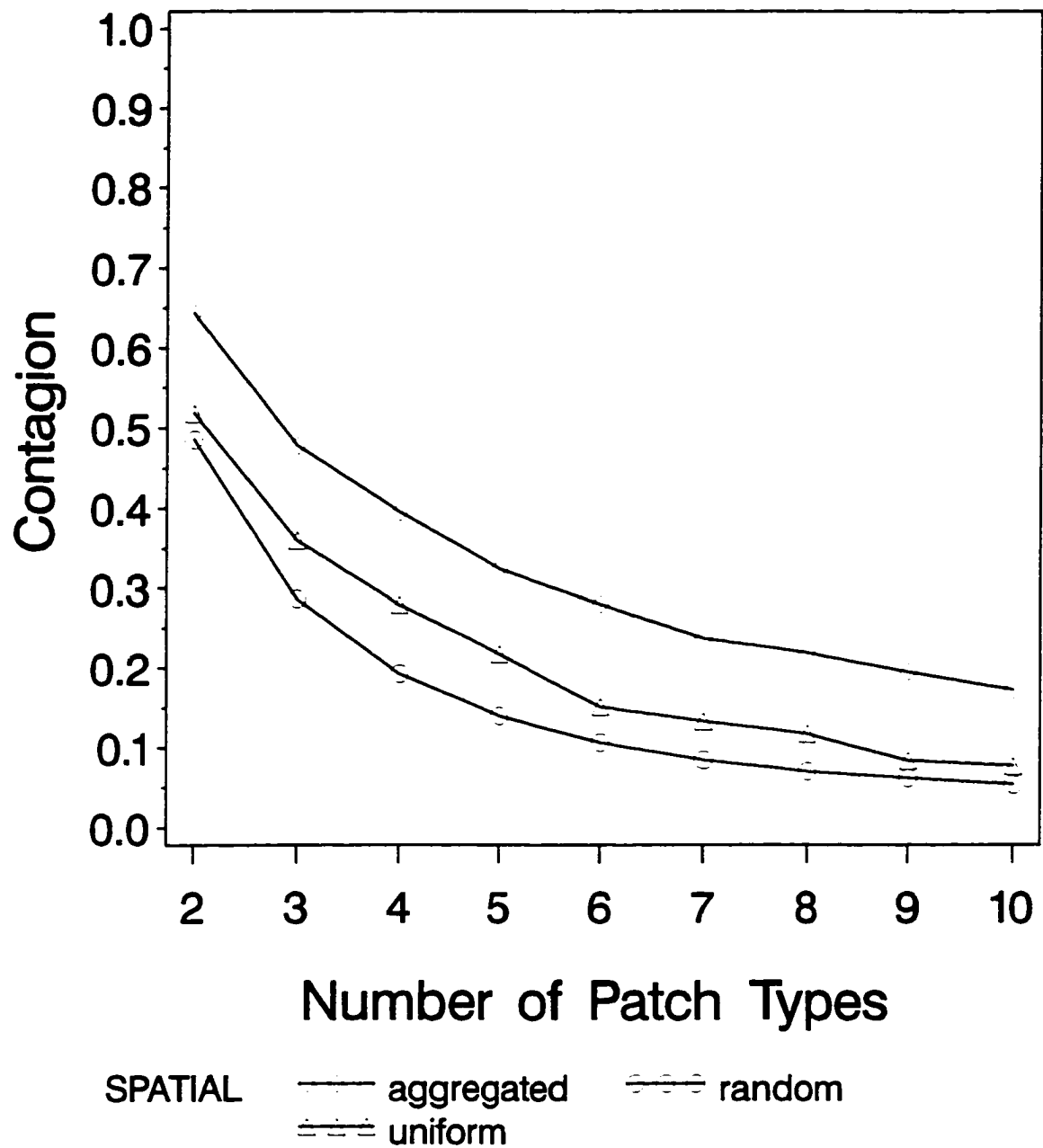


Figure 16. Trend lines showing relationship between Γ_1 and the two controlled variables: spatial pattern and number of patch types. Relative evenness ≈ 1 at each point along the x-axis.

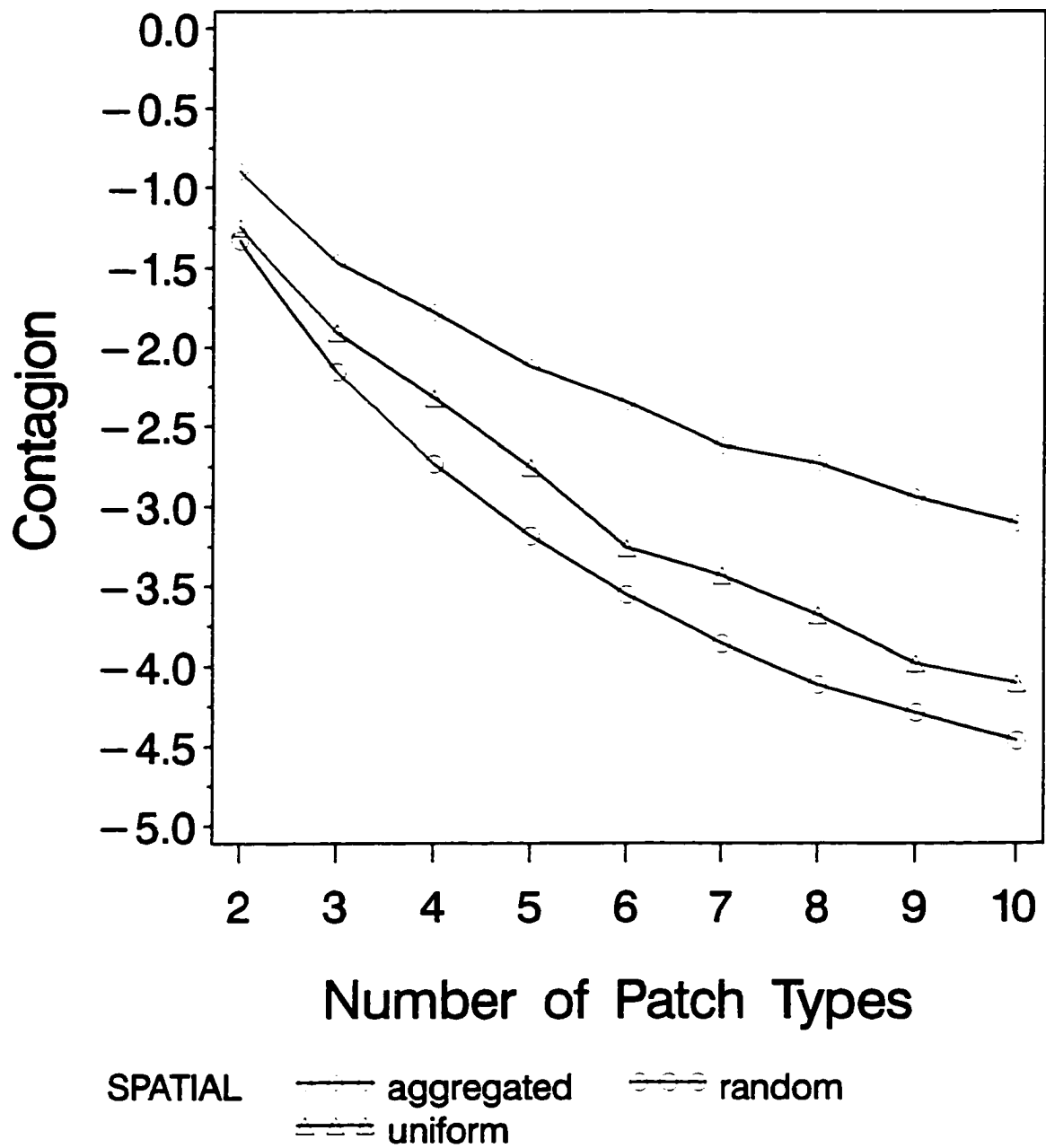


Figure 17. Trend lines showing relationship between Γ_2 and the two controlled variables: spatial pattern and number of patch types. Relative evenness ≈ 1 at each point along the x-axis.

Figure 18 displays a graph of the RC_2 values. As before, the three spatial configurations separate out very cleanly, but the three lines are essentially flat, showing that RC_2 is completely insensitive to changes in number of patch types on the landscape when category proportions are equal. What does this mean?

The Case Against Relative Contagion

Figure 18 reveals that RC_2 is measuring something other than contagion. For each value along the x-axis RE is constant, at its maximum value of 1. There is a direct parallel between the constancy of RE and the constancy of RC_2 for each spatial configuration. It appears RC_2 is measuring evenness. To illustrate this point further, consider the three uniform landscapes in Figure 19. Conceptually, as more patch types are included on a landscape, contagion should decrease. Because patch type proportions on landscape A change only slightly to create landscapes B and C with one and then two additional patch types, there should be a small decrease in contagion from landscape A to B, and then from B to C. Table 3 lists the breakdown of category proportions on the three landscapes and gives the index values computed on these landscapes. Both Γ_1 and Γ_2 decrease slightly from A to B to C, as expected. However, the relative contagion indices increase substantially from A to B to C. The relativized indices are not reflecting changes in contagion, but rather they are reflecting changes in evenness, as shown by RE .

The effect of scaling contagion relative to the maximum contagion possible creates indices with mathematically undesirable qualities. Such indices are overly sensitive to small variation in composition. Sampling error could easily sway the result one direction or the other, so if contagion indices are to be meaningful they should be relatively insensitive to

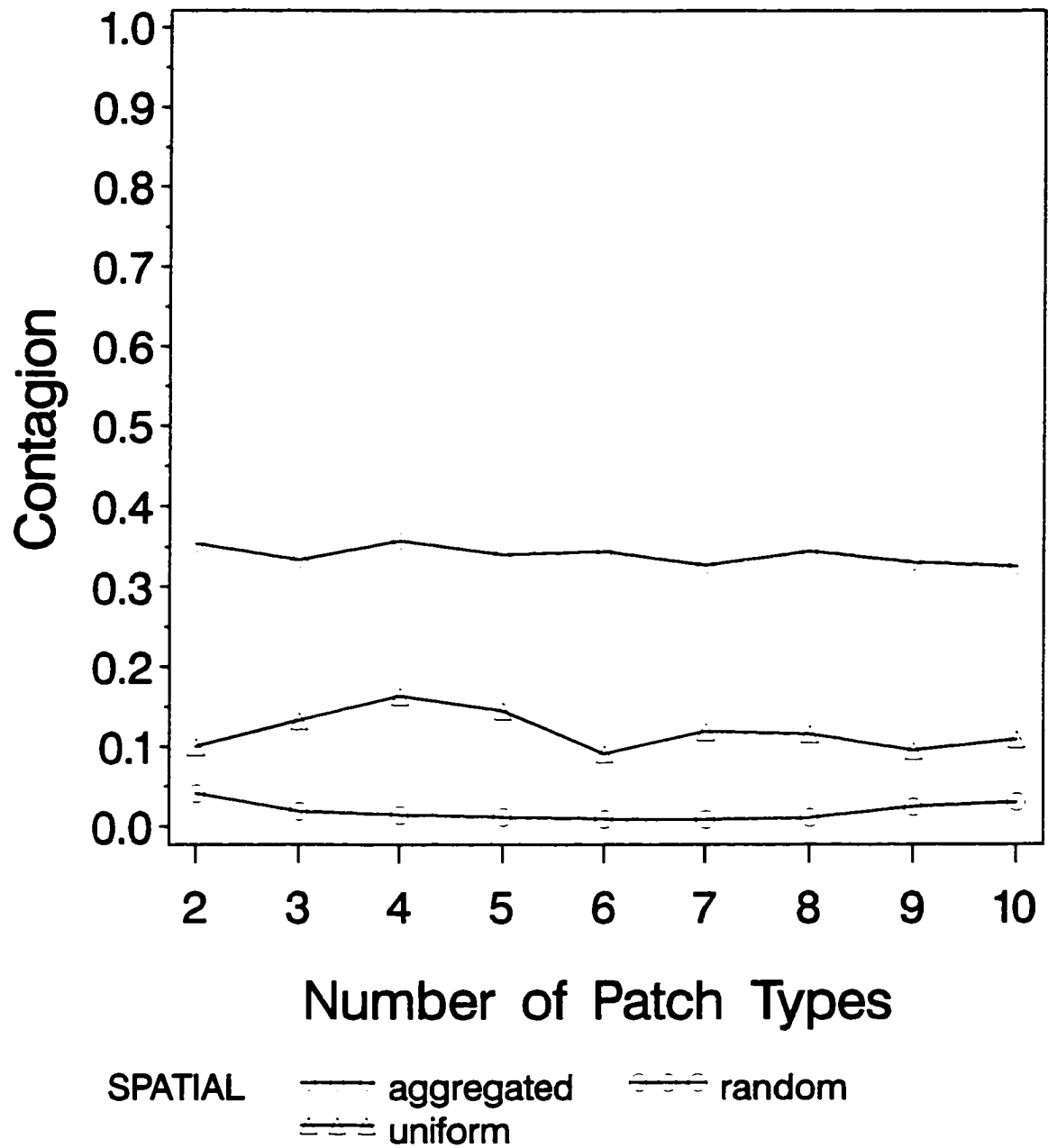


Figure 18. Trend lines showing relationship between RC_2 and the two controlled variables: spatial pattern and number of patch types. Relative evenness ≈ 1 at each point along the x-axis.

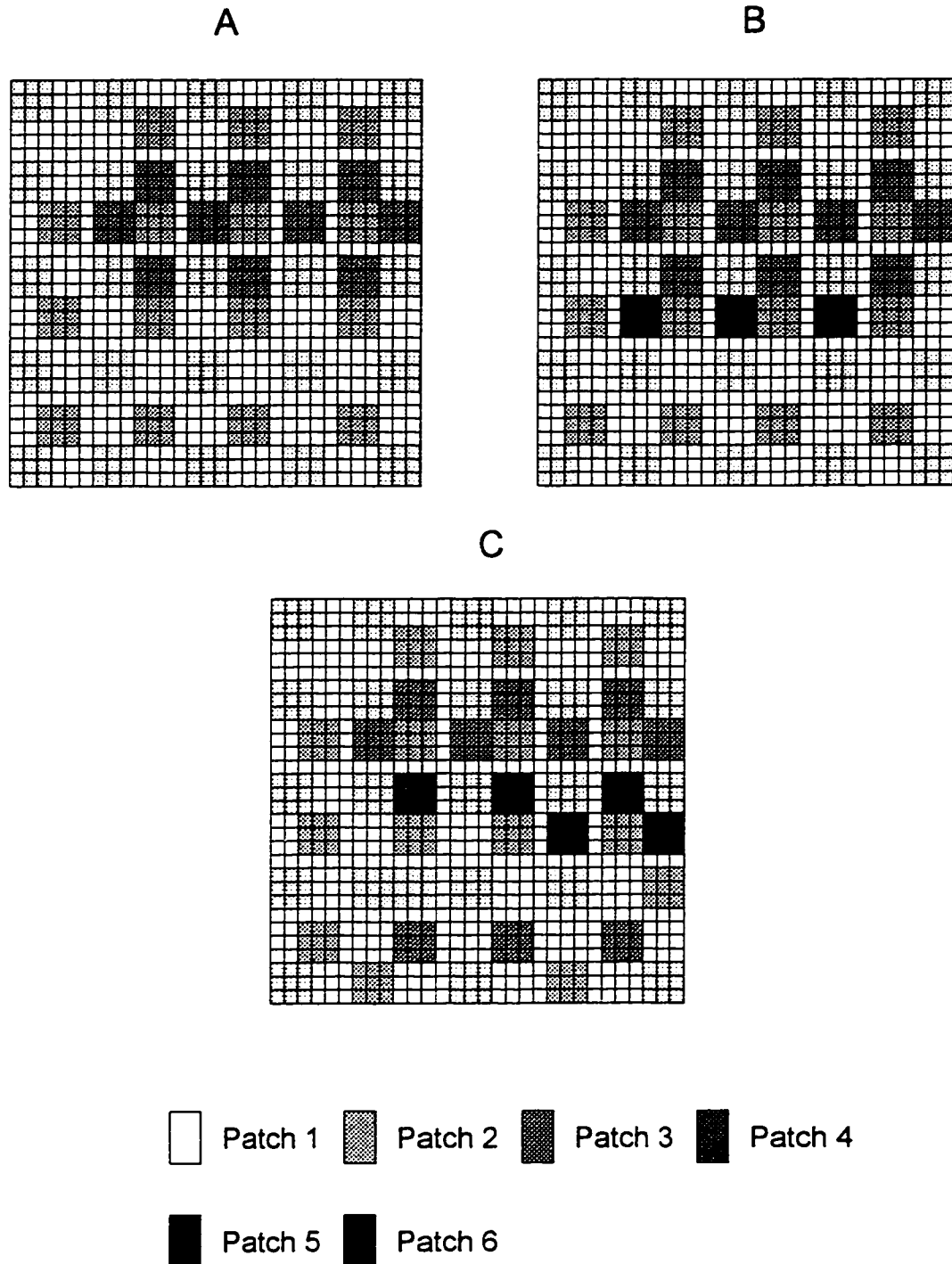


Figure 19. Three uniform landscapes; (A) has 4 patch types, (B) has 5 patch types, and (C) has 6 patch types.

Table 3. Illustrative example showing illogical behavior of relative contagion indices.

Landscape	Patch Type						Index					
	1	2	3	4	5	6	RE	Γ_1	Γ_2	RC_1	RC_2	$RC_{unordered}$
	-----proportion-----						-----value-----					
A	0.50	0.25	0.15	0.10			0.77	0.31	-2.22	0.26	0.20	0.21
B	0.47	0.25	0.15	0.10	0.03		0.71	0.28	-2.41	0.31	0.25	0.27
C	0.47	0.23	0.15	0.10	0.03	0.02	0.66	0.27	-2.52	0.37	0.30	0.32

such a change. This revelation is not new. Sheldon (1969) and Peet (1975) argued against relativizing species diversity indices and they demonstrated that such indices are mathematically untenable.

The inappropriateness of using EE_{\max} in an index of contagion is not restricted to the form of ratios. Let us examine the approach used by O'Neil et al. (1988b). They proposed an index using EE_{\max} as an additive term (see Equation (5)), though they incorrectly specified EE_{\max} . For our purpose, let $C(p_{ij}) = E[X|p_{ij}] \times E[1/(X + 1)|p_{ij}]$ and using this result in Equation (20) with $\phi = EE_{\max} = 2 \ln(n)$ we obtain

$$D_2^* = 2 \ln(n) + \sum_{i=1}^n \sum_{j=1}^n p_{ij} \ln(p_{ij}) \quad (48)$$

which is the same as Equation (5) but with the correct EE_{\max} term. This index, of course, is bounded by 0 and $2 \ln(n)$. I computed D_2^* on the simulated landscapes of Figures 7, 8, and 9, and the values are listed in Appendix E, Table E1. A graph of these 27 values is presented in Figure 20. Illogical behavior is very apparent, with the lines increasing and then decreasing, while each line should monotonically decrease. The moral is that contagion indices, if they are to retain meaning, should not be relativized.

The Meaning of Relative Contagion

In this section I point out the meaning of relative contagion and I examine the relationship between Γ_2 and RC_2 . To begin, the concept of diversity confounds the notions of richness and evenness (Pielou 1977). It is well established in the ecological literature that the evenness or equitability component of diversity can be measured independent from richness in two ways which converge for large sample sizes: evenness = D/D_{\max} and

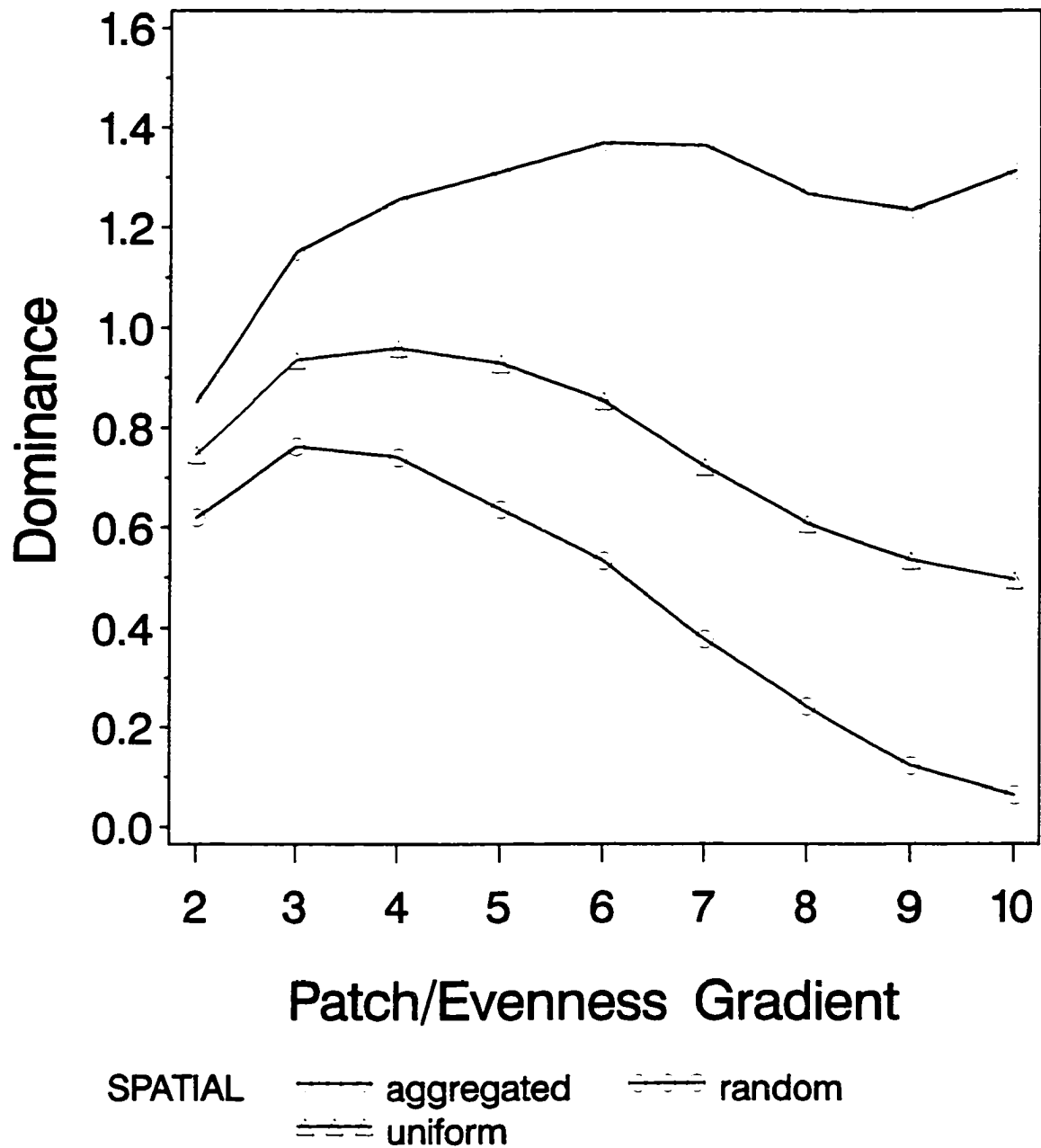


Figure 20. Trend lines showing relationship between D_2^* and the two controlled variables: spatial pattern and number of patch types. There is an increasing gradient of evenness along the x-axis.

evenness = $(D - D_{\min}) / (D_{\max} - D_{\min})$ where D_{\min} and D_{\max} refer to the minimum and maximum value that a diversity index D can attain (See Pielou 1975, pp. 14-17; Magurran 1988, pp. 36-37). Hence an index of evenness is a relativized diversity index. Contagion confounds richness, evenness, and spatial pattern. Relativizing contagion has the effect of creating an evenness index while retaining the configuration component (see Figure 18 and Table 3). Even Li and Reynolds (1993) state: “RC is in essence a function of an evenness index, EE/EE_{\max} ; thus, it has all the advantages and disadvantages of an evenness index.” Apparently they were not fully aware of the mathematical implications of relativizing contagion. Since RC is not a true measure of contagion, and evenness is better quantified using RE , there seems little justification for use of relative contagion.

The relationship between Γ_2 and RC_2 is now clear. Recall that the entropy-based Shannon species diversity index is $H' = \sum_{i=1}^s p_i \ln(p_i)$, and when scaled by its maximum possible value, $\ln(s)$, it becomes an index of evenness, and is no longer an index of diversity. The contagion analogue of the diversity index H' is Γ_2 . Therefore, the true entropy index of contagion is Γ_2 ; and RC_2 , the scaled or relativized version of Γ_2 , is its corresponding “evenness/configuration” index.

Monte Carlo Investigation of Distribution

Figure 21 displays the three 24×24 lattice simulated landscapes used in the Monte Carlo investigation. Theoretical variances for $\hat{\Gamma}_1$ and $\hat{\Gamma}_2$ were computed from Equations (40) and (42) using program CONTAGION. Nonparametric variances for $\hat{\Gamma}_1$ and $\hat{\Gamma}_2$ were computed from Equation (18) using program VAREST. Table 4 gives the results. There is quite close agreement between the theoretical variances and the nonparametric variances

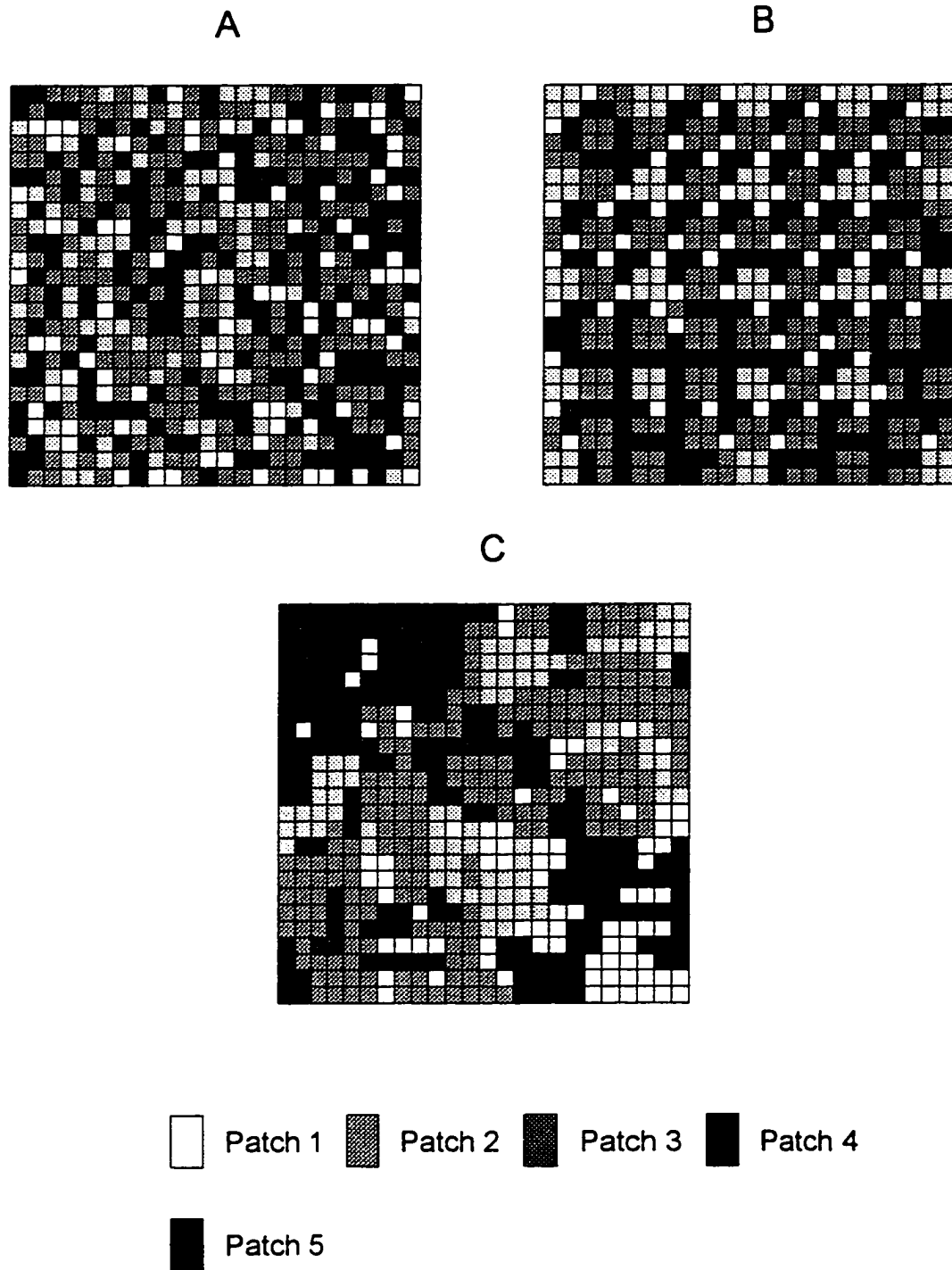


Figure 21. Simulated (A) random, (B) uniform, and (C) aggregated landscapes used in Monte Carlo investigation of contagion index variance and distribution.

Table 4. Comparison of theoretical variance against nonparametric estimate of variance on three simulated 24×24 lattice landscapes of five patch types each.

Random Landscape			
SUBREGION	$\hat{\Gamma}_1$	$\hat{\Gamma}_2$	RE
1	.1679	-3.004	.8901
2	.1604	-3.049	.9254
3	.1652	-3.031	.8913
4	.1728	-2.978	.8906
Nonparametric:	$\text{var}(\hat{\Gamma}_1) = .00002026$	$\text{var}(\hat{\Gamma}_2) = .0007181$	
Theoretical:	$\text{var}(\hat{\Gamma}_1) = .00001243$	$\text{var}(\hat{\Gamma}_2) = .0004573$	
Uniform Landscape			
SUBREGION	$\hat{\Gamma}_1$	$\hat{\Gamma}_2$	RE
1	.1678	-2.993	.9172
2	.1747	-2.954	.8966
3	.1768	-2.956	.8888
4	.1702	-2.985	.8986
Nonparametric:	$\text{var}(\hat{\Gamma}_1) = .00001256$	$\text{var}(\hat{\Gamma}_2) = .0002962$	
Theoretical:	$\text{var}(\hat{\Gamma}_1) = .00001673$	$\text{var}(\hat{\Gamma}_2) = .0005370$	
Aggregated Landscape			
SUBREGION	$\hat{\Gamma}_1$	$\hat{\Gamma}_2$	RE
1	.2662	-2.542	.7241
2	.2790	-2.435	.6394
3	.2782	-2.469	.6360
4	.2565	-2.523	.7478
Nonparametric:	$\text{var}(\hat{\Gamma}_1) = .00008615$	$\text{var}(\hat{\Gamma}_2) = .001800$	
Theoretical:	$\text{var}(\hat{\Gamma}_1) = .00004348$	$\text{var}(\hat{\Gamma}_2) = .0009992$	

in all the cases. The nonparametric approach in general overestimated the variances (in 4 of the 6 cases) as compared to the parametric formulas (Equations (40) and (42)). This outcome was not unexpected because parametric procedures are more powerful and optimum if parametric assumptions are met. Hence I expected the theoretically estimated variances to be smaller than the corresponding nonparametric estimates. The estimates were sufficiently close to substantiate the validity of Equations (40) and (42).

Using program REPHIST, replicate histograms were constructed via Equation (19) for $\hat{\Gamma}_1$ and $\hat{\Gamma}_2$ from the three landscapes of Figure 21. Histogram plots for $\hat{\Gamma}_1$ and $\hat{\Gamma}_2$ are shown in Figures 22 and 23, respectively. Bear in mind that a replicate histogram provides an approximation of the unknown distribution of a statistic, it is not an exact picture of the sampling distribution; as such smoothed histogram plots tend to be wavy or squiggly (see examples in Sherman and Carlstein 1996). In both figures, the histogram plots for landscapes A and B are very smooth, have good symmetry, and are very nearly bell-shaped, though Figure 23B displays some waviness at the peak. Clearly Figure 22A and 22B along with Figure 23A and 23B are representative of normal distributions. Figures 22C and 23C display a wave on the upper tail and show some mild right skewness. However, the overall shape of these two graphs is smooth and clearly bell-like. These graphs do not indicate a major departure from normality. It is reasonable to conclude that all of the histogram plots are compatible with a normal sampling distribution. This validates the asymptotic normality results of the “SAMPLING PROPERTIES” section.

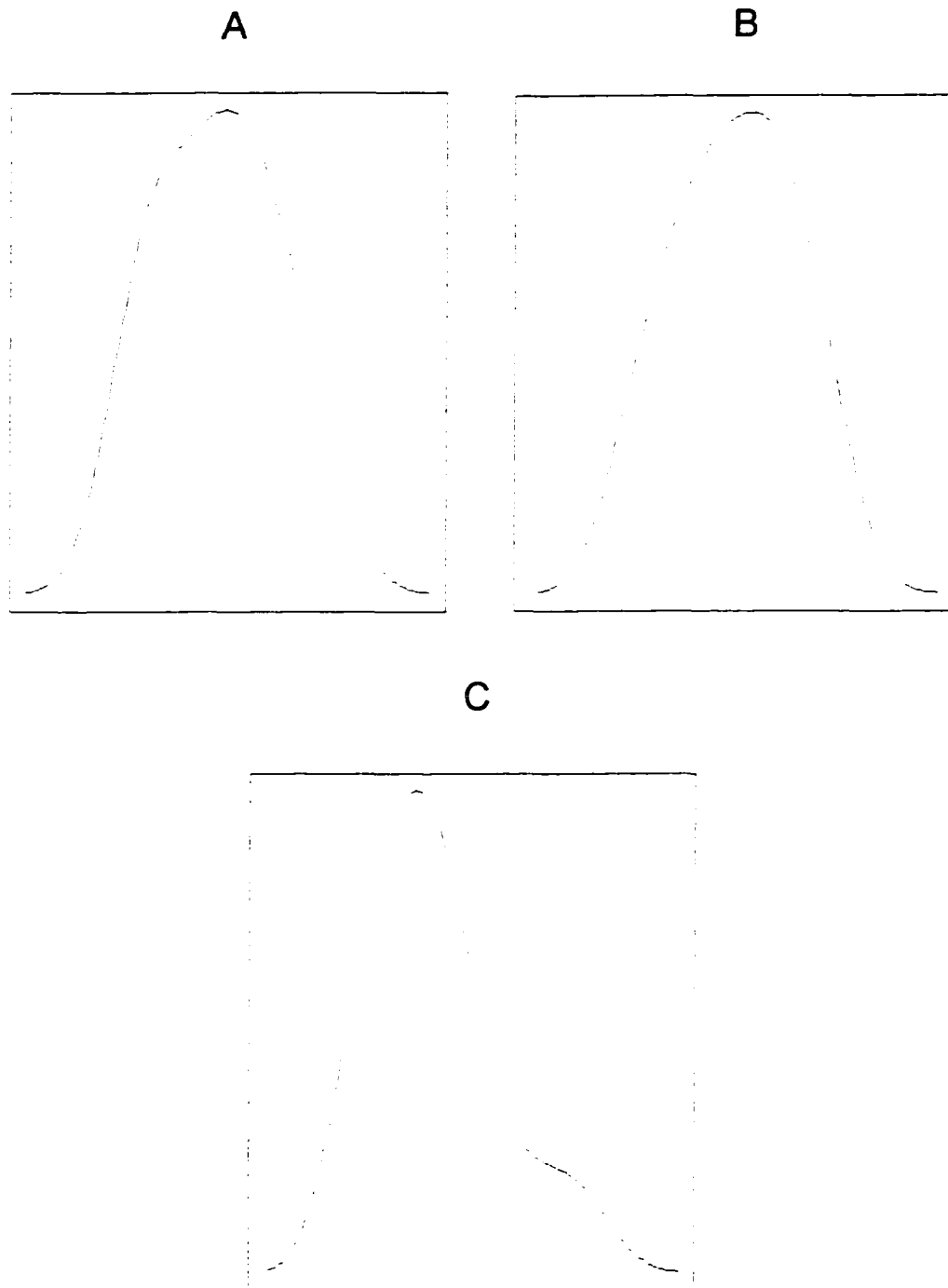


Figure 22. Kernel smoothed replicate histograms of $\hat{\Gamma}_1$ using the (A) random, (B) uniform, and (C) aggregated landscapes of Figure 21.

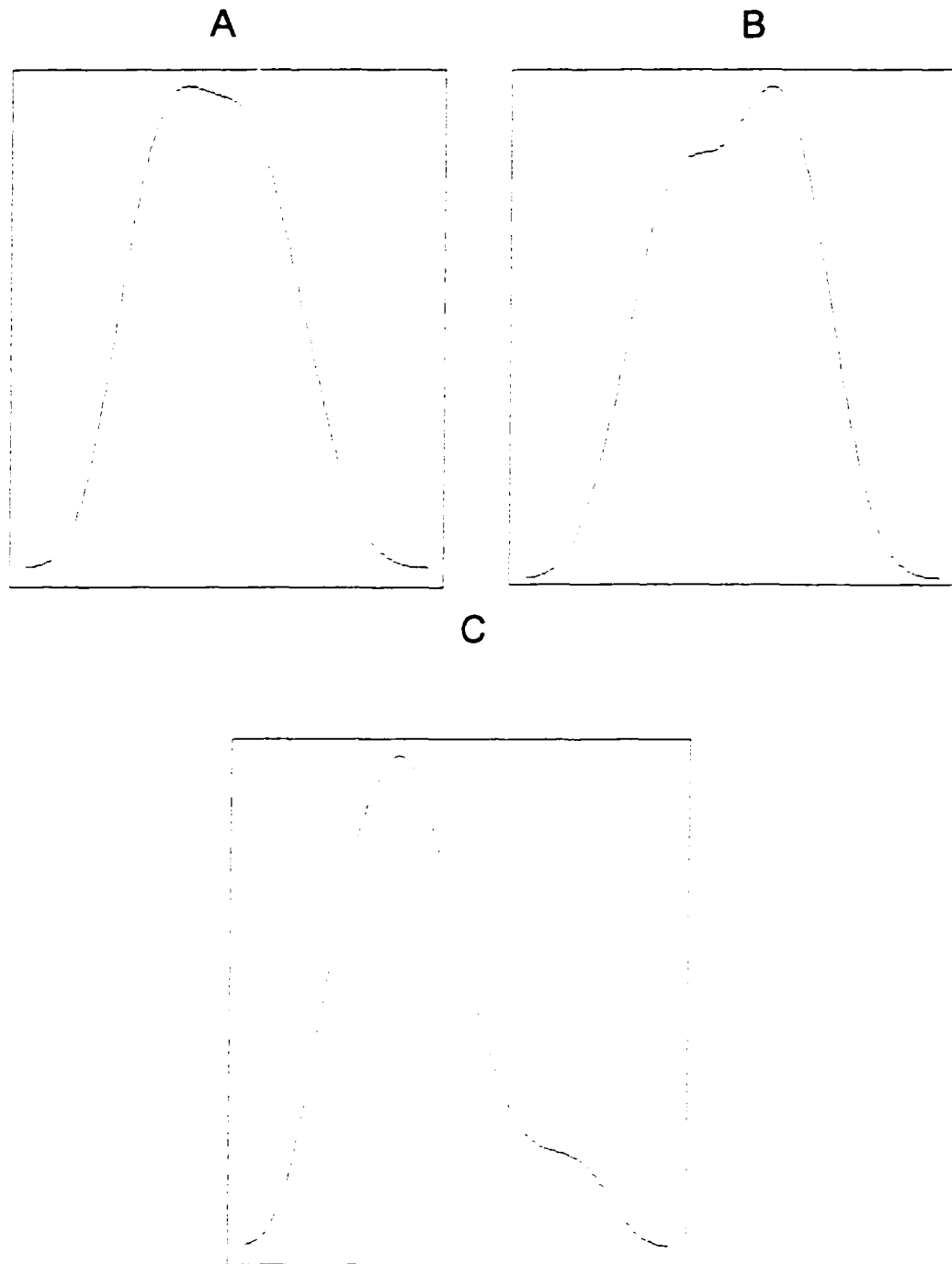


Figure 23. Kernel smoothed replicate histograms of $\hat{\Gamma}_2$ using the (A) random, (B) uniform, and (C) aggregated landscapes of Figure 21.

Analysis of Contagion on the Physiographic Provinces

Table 5 lists the $\hat{\Gamma}_1$ contagion values and their corresponding variances computed on each of the provinces at the three survey periods. All $\hat{\Gamma}_1$ values in Table 5 are fairly low, reflecting the fact that all three landscapes have many patches. Compositional change, in terms of changing proportions of forest cover types (see Table 2), is probably more responsible for the slight differences in $\hat{\Gamma}_1$ values than configuration. From the tests of hypotheses in Table 5 we can see that there was a significant increase in contagion in the Great Appalachian Valley in 1990 over the other two provinces. All in all there is great uniformity in contagion values in Table 5 indicating there has been little or no change in processes affecting contagion over the last two decades and that all three provinces operate under the same influences.

The $\hat{\Gamma}_2$ contagion values and their corresponding variances computed on each of the provinces at the three survey periods are listed in Table 6. This table tells much the same story as Table 5, but there are some interesting differences. The first hypothesis test in Table 6 indicates that contagion has increased significantly in the Great Appalachian Valley in 1990 over the previous two periods. The fourth hypothesis test tells us that contagion is lower in the Great Appalachian Valley than in the other two provinces in 1972. The fifth hypothesis test indicates that contagion is higher in the Piedmont Province than in the other two provinces in 1982. Hence there were three significant tests using $\hat{\Gamma}_2$ whereas there was only one significant test using $\hat{\Gamma}_1$. From this we can conclude that use of $\hat{\Gamma}_1$ provides a more conservative test. Let us examine composition and configuration on the landscapes to try and understand the significant tests.

Table 5. Contagion values, variances, and F -tests using $\hat{\Gamma}_1$ on three physiographic provinces in Alabama during three survey periods.

Province	Year	$\hat{\Gamma}_1$	$\text{var}(\hat{\Gamma}_1)$	T
Great Appalachian Valley	1972	0.1713	.00003415	407
Great Appalachian Valley	1982	0.1798	.00004075	407
Great Appalachian Valley	1990	0.1888	.00003319	407
Blue Ridge-Talladega Mt.	1972	0.1736	.00004874	155
Blue Ridge-Talladega Mt.	1982	0.1725	.00005597	155
Blue Ridge-Talladega Mt.	1990	0.1709	.00005583	155
Piedmont	1972	0.1778	.00002236	409
Piedmont	1982	0.1864	.00001777	409
Piedmont	1990	0.1787	.00001766	409

1) For Appalachian, $H_0: \Gamma_{72}=\Gamma_{82}=\Gamma_{90}$; $F=2.128$, Prob=0.120

2) For Blue Ridge, $H_0: \Gamma_{72}=\Gamma_{82}=\Gamma_{90}$; $F=0.035$, Prob=0.965

3) For Piedmont, $H_0: \Gamma_{72}=\Gamma_{82}=\Gamma_{90}$; $F=1.154$, Prob=0.316

4) For 1972, $H_0: \Gamma_A=\Gamma_B=\Gamma_P$; $F=0.349$, Prob=0.706

5) For 1982, $H_0: \Gamma_A=\Gamma_B=\Gamma_P$; $F=1.599$, Prob=0.203

6) For 1990, $H_0: \Gamma_A=\Gamma_B=\Gamma_P$; $F=2.926$, Prob=0.054

Note: The subscript A stands for the Great Appalachian Valley, B is for the Blue Ridge-Talladega Mountain, and P is for the Piedmont.

Table 6. Contagion values, variances, and F -tests using $\hat{\Gamma}_2$ on three physiographic provinces in Alabama during three survey periods.

Province	Year	$\hat{\Gamma}_2$	$\text{var}(\hat{\Gamma}_2)$	T
Great Appalachian Valley	1972	-3.0609	.001819	407
Great Appalachian Valley	1982	-3.0167	.002158	407
Great Appalachian Valley	1990	-2.9075	.001815	407
Blue Ridge-Talladega Mt.	1972	-2.9513	.002264	155
Blue Ridge-Talladega Mt.	1982	-2.9813	.002895	155
Blue Ridge-Talladega Mt.	1990	-2.9747	.002111	155
Piedmont	1972	-2.9517	.001325	409
Piedmont	1982	-2.8583	.001088	409
Piedmont	1990	-2.9340	.001286	409

1) For Appalachian, $H_0: \Gamma_{72} = \Gamma_{82} = \Gamma_{90}$; $F=3.231$, Prob=0.040

2) For Blue Ridge, $H_0: \Gamma_{72} = \Gamma_{82} = \Gamma_{90}$; $F=0.102$, Prob=0.903

3) For Piedmont, $H_0: \Gamma_{72} = \Gamma_{82} = \Gamma_{90}$; $F=1.997$, Prob=0.136

4) For 1972, $H_0: \Gamma_A = \Gamma_B = \Gamma_P$; $F=2.451$, Prob=0.087

5) For 1982, $H_0: \Gamma_A = \Gamma_B = \Gamma_P$; $F=3.839$, Prob=0.022

6) For 1990, $H_0: \Gamma_A = \Gamma_B = \Gamma_P$; $F=0.779$, Prob=0.459

Note: The subscript A stands for the Great Appalachian Valley, B is for the Blue Ridge-Talladega Mountain, and P is for the Piedmont.

Composition

From Table 2 it is apparent that the longleaf-slash, oak-gum-cypress, and elm-ash-cottonwood types are minor components of the provinces. Keep in mind that proportions in Table 2 are point estimates and they have statistical error. Proportions range from less than 1% for the elm-ash-cottonwood type in the Great Appalachian Valley Province to about 5% for longleaf-slash in the Blue Ridge-Talladega Mountain Province. Though minor, these forest types nonetheless play a role in determining the contagion of the provinces.

Again referring to Table 2, there has been a 4% to 7% decrease in the loblolly-shortleaf type from 1972 to 1990 in the three provinces. The oak-pine type has declined by 4% in the Great Appalachian Valley; it declined by 5% in 1982 but regained the lost 5% in 1990 in the Blue Ridge-Talladega Mountain Province; and it declined nearly 5% in 1982 but regained 3% in 1990 on the Piedmont Province. The oak-hickory type has increased notably in the Great Appalachian Valley (8%) and the Piedmont (11%) Provinces between 1972 and 1990, and was up 7% in 1982 but lost 4% in 1990 in the Blue Ridge-Talladega Mountain Province. The nonforest proportion has been relatively stable, shifting only 1% to 3.5% in the provinces between survey periods. It is worth noting that the Great Appalachian Valley Province has considerably more nonforest area, about 36% to 38%, compared to the other two provinces, which have around 19% to 23% nonforest area.

Configuration

The landscape level views of forest cover types in Figures 4-6 reveal subtle rather than dramatic changes in pattern and composition occurring in the provinces between 1972 and 1990. For the three physiographic provinces the pattern appears to be a mix of random

and clumped. Obviously, all provinces display a degree of fragmentation. Fragmentation at this geographic level results from a complex mix of cities and towns, dams and artificial lakes, forest type conversion, farming, and inherent site conditions such as topography, soil type, and moisture.

The Great Appalachian Valley Province (Figure 4) has a topography of folded mountain ridges with interspersed valleys and hills. The mountain ridges are oriented from southwest to northeast. In Figure 4 we can see that the nonforest type occurs in three concentrated clumps. These clumps are the result of two major cities and a man made lake, with Birmingham in the southwest and Gadsden and Weiss Lake in the northeast. The Coosa River flows through this province coming in from the east. The Coosa was dammed to make Weiss Lake. The Coosa flows southwestward below Gadsden and continues through the province eventually crossing the very western part of the Blue Ridge-Tallega Mountain Province and the Piedmont Province. In the Great Appalachian Valley Province the oak-gum-cypress and the elm-ash-cottonwood types occur in the broad valley areas along the Coosa and its tributaries. The ridges of this province are of either sandstone or shale and chert. The sandstone ridges are overlain by silt loams of fair to good productivity for both pines and hardwoods. The shale soils are shallow to deep and are of poor to good productivity for forests. The chert ridge soils are deep soils of medium to good forest quality. The ridges and slopes support extensive areas of hardwood forests and we can see clumping of the oak-hickory type in Figure 4. Pine management, both planted and natural, is important in this region and we can see some clumping of the loblolly-shortleaf type. The other types are more scattered.

The smallest of the three provinces, the Blue Ridge-Talladega Mountain (Figure 5) is a topographically reduced extension of the heavy slates and quartzites of the western Blue Ridge Mountains of East Tennessee and North Georgia. The high mountain ridges have shallow, stony residual soils but the slopes have colluvial soils of medium to good productivity for hardwoods and pines. The lesser mountain ridges and hills of moderate to steep slopes have silt loam to loam topsoils of medium to very good site quality for pines and fair to good site quality for hardwoods. This province lacks the elm-ash-cottonwood type and the oak-gum-cypress type (except in 1982 where it shows up at a plot near where the Coosa crosses the province). The nonforest areas are scattered but one clump shows up at the northeast neck of the province coinciding with the city of Anniston. Immediately to the west of the “neck” is a narrow cut out area belonging to the Great Appalachian Valley which contains a piece of this nonforest city area. We can see in Figure 5 that this province has some clumping in the loblolly-shortleaf and oak-hickory types which relates to the topography and soil characteristics of the province.

Heading south out of the Blue Ridge-Talladega Mountain Province we drop in elevation and enter into the Piedmont Province (Figure 6). Cultivation was once widespread in this province. Topography can be rugged in the schist ridges, which lie parallel to the ridges of the Blue Ridge-Talladega Mountain Province to the northwest. The granitic hills, south of the schist ridges, are rounded and worn, and are considered fair to good for growing loblolly, shortleaf, and longleaf pines. The topography turns to gently rolling plains and plateaus to the southeast. The Tallapoosa River generally runs north to south nearly through the center of the province with a major man-made lake, Lake Martin, formed close

to the southern border. In Figure 6 this lake shows up as a large nonforest block near the southern border a little to the west of center. In the southeast of the province there is a northeast to southwest line of towns (Lanett and Valley, Opelika, and Auburn) along interstate 85 which show up as three small nonforest blocks. Farming, though not as extensive as in years past, is still important in the region, creating small nonforest fragments across the province. Again we can see in Figure 6 that there is clumping of the loblolly-shortleaf and oak-hickory types.

Remarks

It is recognized that FIA data have low resolution for analysis of forest cover types on landscapes. More than one cover type boundary could be crossed (and consequently undetected) when moving from a lattice plot to a nearest neighbor plot. This problem of missing connections is inherent in all join-count based statistics regardless of the type of categorical data. This, however, does not negate their use. Different sampling realizations will determine a different connection matrix. Certainly the distance d_{ij} between points i and j play a crucial role in the ability to detect the spatial autocorrelation present. I believe broad spatial patterns can nonetheless be discerned with FIA data.

From my examination of composition and configuration and the contagion tests, I can generally conclude that (i) there has been a minor shift in loblolly-shortleaf and oak-pine acreage into oak-hickory acreage, (ii) the overall landscape pattern in the three provinces is similar being a mix of random and clumped, (iii) the loblolly-shortleaf and oak-hickory forest cover types occur in moderate to large clumps while the others occur in small (relative to the landscape) well intermixed patches, and (iv) there have been only small,

though in some cases statistically significant, changes in contagion on the provinces. Both $\hat{\Gamma}_1$ (Table 5, test 6) and $\hat{\Gamma}_2$ (Table 6, test 1) indicate a significant increase in contagion for the Great Appalachian Valley in 1990. The large proportion of nonforest and clumpiness of this type are probably most responsible for this result. In 1972 contagion appears to be lowest on the Great Appalachian Valley Province (Table 6, test 4). This province has the greater richness which may account for its lower contagion. Finally, in 1982 the Piedmont has greater contagion over the other two provinces (Table 6, test 5). The Piedmont experienced greater composition changes from 1972 to 1982 than the other provinces which probably accounts for its greater contagion.

A host of tools are today available to policy makers and forestry practitioners to assess the state of forests and to help guide in their management for sustainability. In dealing with landscape or regional scales of resolution, GIS and landscape indices are important tools for characterizing and comparing landscape diversity.

Choice of Contagion Index

Based on the simulations, both Γ_1 and Γ_2 behaved in a manner consistent with the concept of contagion. Both indices were sensitive to changes in composition and configuration and correctly ordered spatial pattern with aggregated > uniform > random. The Γ_1 values displayed more curvature (see Figures 10 and 16) than the Γ_2 values (see Figures 11 and 17), hence Γ_2 gave better separation between index values when number of patch types exceeded seven. The disadvantage of Γ_2 is its negative scale. One could specify a positive constant for ϕ (see Equation (20)) and thereby translate the axis to give positive readings, but there is no constant other than $\phi = 0$ that will retain the origin as an end point

in the range interval. We have already seen the deleterious effects of specifying a variable value for ϕ such as $2 \ln(n)$.

From a theoretical point of view there is nothing inherently wrong with having a negative scale, but from a practical point of view most users of indices prefer a positive scale. Thus the index Γ_1 is appealing because of its positive scale. Also, Γ_1 possesses the fixed range $(0, 1]$ providing a ready interpretation of no contagion at the lower extreme and perfect contagion at the upper extreme, with degrees of contagion in between. The index Γ_2 has a nonstationary range dependent on the quantity $2 \ln(n)$, which adds a layer of complexity in interpreting contagion on the range of Γ_2 .

From the analysis of contagion on the physiographic provinces, use of $\hat{\Gamma}_1$ in the analysis of variance appears to give a more conservative test of contagion. Because $\hat{\Gamma}_2$ has an expanding range with increasing number of patch types, one could raise the issue of what is a meaningful difference. Past a certain point of fragmentation or with increasing evenness and number of categories, $\hat{\Gamma}_1$ tends to flatten out whereas $\hat{\Gamma}_2$ keeps increasing its range and can take on lower and lower values. Authors such as O'Neil et al. (1988b), Riitters et al. (1996), and myself, want measures of contagion to assess dominance or concentration². Philosophically then, I believe that below a certain point differences have little functional significance on the landscape. The "conservativeness" of $\hat{\Gamma}_1$ reflects this philosophical niche. If one does not subscribe to this ideology, then $\hat{\Gamma}_2$ would seem preferable.

²In ecology, dominance or concentration is normally measured using some function of p_i^2 as this has been shown to reflect expected commonness (see Pielou 1975, pp. 8-9). Note that Γ_1 is a function of p_{ij}^2 .

Both indices Γ_1 and Γ_2 provide acceptable measures of contagion. Relative contagion indices, as previously demonstrated, do not provide acceptable measures of contagion. The choice between Γ_1 and Γ_2 is largely one of ideology, with Γ_1 having a more “user friendly” fixed positive range $(0,1]$ as opposed to the $[-2\ln(n),0]$ range of Γ_2 .

SUMMARY AND CONCLUSIONS

A quantitative basis for measuring spatial structure and composition is the contagion index. Landscape contagion, as a measure, has a reverse scale from that of species diversity. Thus, higher values of contagion result from landscapes with a few large, contiguous patches, whereas lower values characterize landscapes with many small patches.

Only a few contagion indices exist; and most, if not all, have been scaled relative to their maximum value. Landscape indices have been used primarily in quantifying remotely sensed images. Field-determined forest cover type data are available from the USDA Forest Service, Forest Inventory and Analysis (FIA) surveys from sampling lattices covering the whole of the U.S. Some inherent problems with the FIA data are low resolution, plots are spaced three miles apart, and all nonforest land are classified the same.

Sampling properties have previously been worked out for the simple binary join-count statistic. Sampling properties are needed for the more complex contagion index based on multi-join counts.

In a contagious landscape the typical patch type is relatively concentrated. Therefore, a generalized measure of contagion can be constructed as a function of concentration. Viewing join counts as a general scheme in waiting times, geometric random variables result. Expected values of the random variables cast into the generalized measure resulted in new contagion indices, which I called Γ_1 and Γ_2 . A widely used relative contagion index, RC_2 , is simply a scaled or relativized version of Γ_2 .

A contagion index can be considered as a sum of multinomial probabilities. Using properties of the multinomial distribution in a Taylor series expansion of $\hat{\Gamma}_1$, $\hat{\Gamma}_2$, and

$\hat{\Gamma}_{2(\text{scaled})}$ (i.e., RC_2), the bias of these statistics is shown to be $o(T^{-1})$ where T is total number of observations. Using the delta method, a standard linear approximation formula, variance equations were derived for $\hat{\Gamma}_1$, $\hat{\Gamma}_2$, and $\hat{\Gamma}_{2(\text{scaled})}$. A function of a multivariate normal is itself normally distributed. Because multinomial probabilities have a limiting multivariate normal distribution, $\hat{\Gamma}_1$, $\hat{\Gamma}_2$, and $\hat{\Gamma}_{2(\text{scaled})}$ are asymptotically normal.

Using simulated random, uniform, and aggregated landscapes, and increasing numbers of patch types (from 2 to 10) across an increasing gradient of evenness, Γ_1 , Γ_2 , and RC_2 distinguished the different spatial configurations and changing composition. Using simulated random, uniform, and aggregated landscapes, and increasing numbers of patch types (from 2 to 10), but with equal patch type proportions, Γ_1 and Γ_2 still distinguished the spatial configurations and changing composition. However, RC_2 did not distinguish among the changing composition. Further analysis confirmed that relativized indices are mathematically untenable. The index RC_2 is really a function of evenness and the true entropy measure of contagion is Γ_2 . I concluded that there is little justification for using relative contagion.

Subseries analysis and replicate histograms are nonparametric techniques for estimating variance and sampling distribution shape. These techniques provided a validity check of the theoretical variances and distributions of $\hat{\Gamma}_1$ and $\hat{\Gamma}_2$.

An analysis of forest cover type contagion was done on three physiographic provinces in Alabama; the Great Appalachian Valley Province, the Blue Ridge-Talladega Mountain Province, and the Piedmont Province. Data were available for the years 1972, 1982, and 1990. It was generally concluded that (i) there was a minor shift in loblolly-

shortleaf and oak-pine acreage into oak-hickory acreage between 1972 and 1990, (ii) the overall landscape pattern in the three provinces was similar, being a mix of random and clumped, (iii) the loblolly-shortleaf and oak-hickory forest cover types occurred in moderate to large clumps while the others occurred in small (relative to the landscape) well intermixed patches, and (iv) there have been only small, though in some cases statistically significant, changes in contagion on the provinces. For the most part there was great uniformity in contagion values indicating there has been little or no change in processes affecting contagion over the last two decades and that all three provinces operate under the same influences.

Contagion indices are normally computed over a regular (square or rectangular) grid. Joins are based on the four neighboring quadrats or pixels. Since survey plots represent Thiessen polygonal areas (because each county starts a new randomly placed grid), this application is a departure from the normal procedure. No effort was made to weight the count of joins based on join length or other suitable procedures. This is a refinement that should be looked into further.

Landscape diversity is inexorably linked to geographic information. Geographic information systems (GIS) are today an important tool for assessment, management, and monitoring. Layers of information can be combined in various ways to get a picture of the whole, or a visual perspective on change, or to highlight critical features, etc. For this study geographic information, utilizing a GIS, was exploited to highlight regional features (physiographic provinces), landscape features (sample points and forest cover types), and for assessment.

The new indices Γ_1 and Γ_2 provide acceptable measures of contagion. Some people may be uncomfortable with the negative number scale of Γ_2 , hence they would find Γ_1 appealing because of its positive number scale. For most, the choice between Γ_1 and Γ_2 will depend on ideology concerning ecological dominance or concentration.

REFERENCES

- Baskent, E.Z., and G.A. Jordan. 1995. Characterizing spatial structure of forest landscapes. *Canadian Journal of Forest Research* 25: 1830-1849.
- Bettinger, P., G.A. Bradshaw, and G.W. Weaver. 1996. Effects of geographic information systems vector-raster-vector data conversion on landscape indices. *Canadian Journal of Forest Research* 26: 1416-1425.
- Bishop, Y.M.M., S.E. Fienberg, and P.W. Holland. 1975. Discrete multivariate analysis: theory and practice. MIT Press, Cambridge, Mass. 557 p.
- Burgess, R.L., and D.M. Sharpe, eds. 1981. Forest island dynamics in man-dominated landscapes. Springer-Verlag, New York. 310 p.
- Carlstein, E. 1986a. The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The Annals of Statistics* 14(3): 1171-1179.
- Carlstein, E. 1986b. Asymptotic normality for a general statistic from a stationary sequence. *The Annals of Probability* 14(4): 1371-1379.
- Carlstein, E. 1988. Law of large numbers for the subseries values of a statistic from a stationary sequence. *Statistics* 19: 295-299.
- Cihlar, J., L. Hung, and X. QingHan. 1996. Land cover classification with AVHRR multichannel composites in northern environments. *Remote Sensing of Environment* 58(1): 36-51.
- Cramér, H. 1946. Mathematical methods of statistics. Princeton University Press, Princeton, NJ. 575 p.
- Davis, F.W. 1993. Introduction to spatial statistics. p. 17-26 *in* Lecture Notes in Biomathematics Vol. 96: Patch Dynamics. S.A. Levin, T.M. Powell, and J.H. Steele (eds.). Springer-Verlag, New York. 307 p.
- De Cola, L. 1989. Fractal analysis of a classified landsat scene. *Photogrammetric Engineering and Remote Sensing* 55: 601-610.
- Efron, B., and R.J. Tibshirani. 1993. An introduction to the bootstrap. Chapman & Hall, New York. 436 p.
- Environmental Systems Research Institute. 1992. ARC/INFO® version 6 user's guide: ARC command references. Environmental Systems Research Institute, Inc., Redlands, CA.

- Evans, D.L., and R.L. Czaplewski. 1996. Investigation of possible contributions NDVI's have to misclassification in AVHRR data analysis. p. 419-424 *in* Spatial Accuracy Assessment in Natural Resources and Environmental Sciences: Second International Symposium. May 21-23, 1996. H.T. Todd, R.L. Czaplewski, and R.H. Hamre (tech. coords.). General Technical Report RM-GTR-277. U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO. 728 p.
- Farmer, J.D., E. Ott, and J.A. Yorke. 1983. The dimension of chaotic attractors. *Physica* 7D: 153-180.
- Feller, W. 1968. An introduction to probability theory and its applications, volume 1, 3rd ed. Sec. II.7: Examples for waiting times, p. 47-50. John Wiley & Sons, New York. 509 p.
- Gonzalez, R.C., and P.A. Wintz. 1987. Digital image processing, 2nd ed. Addison-Wesley, Reading, MA. 503 p.
- Gray, H.L., and W.R. Schucany. 1972. The generalized jackknife statistic. Marcel Dekker, New York. 308 p.
- Hill, M.O. 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology* 54: 427-431.
- Hodgkins, E.J. M.S. Golden, and W.F. Miller. 1979. Forest habitat regions and types on a photomorphic-physiographic basis: a guide to forest site classification in Alabama-Mississippi. Southern Cooperative Series No. 210. Alabama Agricultural Experiment Station, Auburn University, Auburn, AL. 64 p.
- Hutcheson, K. 1970. A test for comparing diversities based on the Shannon Formula. *Journal of Theoretical Biology* 29: 151-154.
- Judge, G.G., R.C. Hill, W.E. Griffiths, H. Lütkepohl, and T. Lee. 1988. Introduction to the theory and practice of econometrics, 2nd ed. John Wiley & Sons, New York. 1024 p.
- Kempthorne, O., and L. Folks. 1971. Probability, statistics, and data analysis. The Iowa State University Press, Ames, IA. 555 p.
- Krummel, J.R., R.H. Gardner, G. Sugihara, R.V. O'Neill, and P.R. Coleman. 1987. Landscape patterns in a disturbed environment. *Oikos* 48: 321-324.
- Kullback, S.M., M. Kupperman, and H.H. Ku. 1962. Tests for contingency tables and Markov chains. *Technometrics* 4: 573-608.

- Li, H., and J.F. Reynolds. 1993. A new contagion index to quantify spatial patterns of landscapes. *Landscape Ecology* 8: 155-162.
- Li, H., J.F. Franklin, F.J. Swanson, and T.A. Spies. 1993. Developing alternative forest cutting patterns: a simulation approach. *Landscape Ecology* 8(1): 63-75.
- Li, H., and J.F. Reynolds. 1994. A simulation experiment to quantify spatial heterogeneity in categorical maps. *Ecology* 75(8): 2446-2455.
- Li, H., and J.F. Reynolds. 1996. Modeling effects of spatial pattern, drought, and grazing on rates of rangeland degradation: a combined Markov and cellular automaton approach. p. 211-230 *in* Scale in Remote Sensing and GIS. D.A. Quattrochi and M.F. Goodchild (eds.). Lewis Publishers, Boca Raton, FL. 480 p.
- Loehle, C., and G. Wein. 1992. Landscape habitat diversity: an information theoretic measure. Contractor Report to Environmental Research Division, Argonne National Laboratory, Argonne, IL, under contract W-31-109-ENG-38. 22 p.
- Magnussen, S. 1997. Calibrating photo-interpreted forest cover types and relative species compositions to their ground expectations. *Canadian Journal of Forest Research* 27: 491-500.
- Magurran, A.E. 1988. Ecological diversity and its measurement. Princeton University Press, Princeton, NJ. 179 p.
- May, D.M. 1990. Stocking, forest type, and stand size class--the Southern Forest Inventory and Analysis unit's calculation of three important stand descriptors. General Technical Report SO-77. U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station, New Orleans. 7 p.
- McGarigal, K., and B.J. Marks. 1995. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. General Technical Report PNW-GTR-351. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR. 122 p.
- Moran, P.A.P. 1948. The interpretation of statistical maps. *Journal of the Royal Statistical Society, series B* 10: 243-251.
- Naveh, Z., and A.S. Lieberman. 1994. Landscape ecology: theory and application, 2nd ed. Springer-Verlag, New York. 360 p.
- O'Neill, R.V., B.T. Milne, M.G. Turner, and R.H. Gardner. 1988a. Resource utilization scales and landscape pattern. *Landscape Ecology* 2: 63-69.

- O'Neill, R.V., J.R. Krummel, R.H. Gardner, G. Sugihara, B. Jackson, D.L. DeAngelis, B.T. Milne, M.G. Turner, B. Zygmunt, S.W. Christensen, V.H. Dale, and R.L. Graham. 1988b. Indices of landscape pattern. *Landscape Ecology* 1: 153-162.
- Patil, G.P.; Taillie, C. 1982. Diversity as a concept and its measurement. *Journal of the American Statistical Association*. 77(379): 548-561.
- Peet, R.K. 1974. The measurement of species diversity. *Annual Review of Ecology and Systematics* 5: 285-307.
- Peet, R.K. 1975. Relative diversity indices. *Ecology* 56: 496-498.
- Pielou, E.C. 1975. *Ecological diversity*. John Wiley & Sons, New York. 165 p.
- Pielou, E.C. 1977. *Mathematical ecology*. John Wiley & Sons, New York. 385 p.
- Politis, D.N., and J.P. Romano. 1993. On the sample variance of linear statistics derived from mixing sequences. *Stochastic Processes and their Applications* 45: 155-167.
- Politis, D.N., and J.P. Romano. 1994. Large sample confidence regions based on subsamples under minimal assumptions. *The Annals of Statistics* 22(4): 2031-2050.
- Possolo, A. 1991. Subsampling a random field. p. 286-294 *in* *Spatial Statistics and Imaging*, IMS Lecture Notes Monograph Series, Vol. 20. A. Possolo (ed.). Institute of Mathematical Statistics, Haywood, CA. 426 p.
- Rao, C.R. 1965. *Linear statistical inference and its applications*. John Wiley & Sons, New York. 522 p.
- Ratnaparkhi, M.V. 1985. Multinomial distributions. p. 659-665 *in* *Encyclopedia of Statistical Sciences*, Vol. 5. S. Kotz, N.L. Johnson, and C.B. Read (eds.). John Wiley & Sons, New York. 741 p.
- Rényi, A. 1961. On measures of entropy and information. *Berkeley Symposium on Mathematical Statistics and Probability*. 4: 547-561.
- Riitters, K.H., R.V. O'Neil, J.D. Wickham, and K.B. Jones. 1996. A note on contagion indices for landscape analysis. *Landscape Ecology* 11(4): 197-202.
- Ripley, B.D. 1981. *Spatial statistics*. John Wiley & Sons, New York. 252 p.
- Salas, S.L., and E. Hille. 1974. *Calculus: one and several variables*, 2nd ed. John Wiley & Sons, New York. 829 p. + Appendix.

- Salwasser, H. 1990. Conserving biological diversity: a perspective on scope and approaches. *Forest Ecology and Management* 35: 79-90.
- Shannon, C.E., and W. Weaver. 1949. *The mathematical theory of communication*. University of Illinois Press, Urbana, IL. 117 p.
- Sheldon, A.L. 1969. Equitability indices: dependence on the species count. *Ecology* 50: 466-467.
- Shenton, L.R., and K. Hutcheson. 1969. Moments of moments and frequencies: expectation of power-sums related to categorized data. Technical Reprint Series No. 17. University of Georgia Computer Center, Athens, GA. p. 7-39.
- Sherman, M., and E. Carlstein. 1994. Nonparametric estimation of the moments of a general statistic computed from spatial data. *Journal of the American Statistical Association* 89(426): 496-500.
- Sherman, M., and E. Carlstein. 1996. Replicate histograms. *Journal of the American Statistical Association* 91(434): 566-576.
- Society of American Foresters. 1980. *Forest cover types of North America (exclusive of Mexico)*. Society of American Foresters, Washington, DC. 148 p.
- Turner, M.G. 1989. Landscape ecology: the effect of pattern on process. *Annual Review of Ecology and Systematics* 20: 171-197.
- Turner, M.G. 1990. Spatial and temporal analysis of landscape pattern. *Landscape Ecology* 4: 21-30.
- Turner, M.G., G.J. Arthaud, R.T. Engstrom, S.J. Hejl, J. Liu, S. Loeb, and K. McKelvey. 1995. Usefulness of spatially explicit population models in land management. *Ecological Applications* 5(1): 12-16.
- U.S. Department of Agriculture, Forest Service. 1972. *Forest survey handbook*. in *Forest Service Handbook 1809.11. Amendment 6. Section 21.5*. U.S. Department of Agriculture, Washington, DC.
- Wenger, K.F., ed. 1984. *Forestry handbook*, 2nd ed. Sec. 1: *Forest Ecology*, p. 1-63. John Wiley & Sons, New York. (Society of American Foresters, Washington, DC). 1335 P.
- White, H. 1984. *Asymptotic theory for econometricians*. Academic Press, Orlando, FL. 228 p.

APPENDIX A

DERIVATIONS FOR $E[X | p_{ij}]$, $E[1/(X+1) | p_{ij}]$, $E[X/(X+1) | p_{ij}]$, AND BOUNDS FOR EQUATION (23)

The contagion formulas for Γ_1 (Equation (23)) and Γ_2 (Equation (24)) come from particular expected values based on a geometric random variable (see Equation (21)). Let $q_{ij} = (1 - p_{ij})$, the required expected values are:

$$\begin{aligned}
 E[X|p_{ij}] &= \sum_{x=0}^{\infty} x p_{ij} q_{ij}^x \\
 &= p_{ij} q_{ij} + 2p_{ij} q_{ij}^2 + 3p_{ij} q_{ij}^3 + \dots \\
 &= p_{ij} q_{ij} (1 + 2q_{ij} + 3q_{ij}^2 + \dots) \\
 &= p_{ij} q_{ij} (1 - q_{ij})^{-2} = (1 - p_{ij})/p_{ij}
 \end{aligned} \tag{A1}$$

$$\begin{aligned}
 E[1/(X+1)|p_{ij}] &= \sum_{x=0}^{\infty} \frac{1}{x+1} p_{ij} q_{ij}^x \\
 &= p_{ij} + \frac{1}{2} p_{ij} q_{ij} + \frac{1}{3} p_{ij} q_{ij}^2 + \frac{1}{4} p_{ij} q_{ij}^3 + \dots \\
 &= p_{ij} (1 + \frac{1}{2} q_{ij} + \frac{1}{3} q_{ij}^2 + \frac{1}{4} q_{ij}^3 + \dots) \\
 &= p_{ij} (q_{ij} + \frac{1}{2} q_{ij}^2 + \frac{1}{3} q_{ij}^3 + \frac{1}{4} q_{ij}^4 + \dots) / q_{ij} \\
 &= p_{ij} [-\ln(1 - q_{ij})] / q_{ij} \\
 &= -p_{ij} \ln(p_{ij}) / (1 - p_{ij})
 \end{aligned} \tag{A2}$$

$$\begin{aligned}
 E[X/(X+1)|p_{ij}] &= \sum_{x=0}^{\infty} \frac{x}{x+1} p_{ij} q_{ij}^x \\
 &= \frac{1}{2} p_{ij} q_{ij} + \frac{2}{3} p_{ij} q_{ij}^2 + \frac{3}{4} p_{ij} q_{ij}^3 + \dots \\
 &= p_{ij} (\frac{1}{2} q_{ij}^2 + \frac{2}{3} q_{ij}^3 + \frac{3}{4} q_{ij}^4 + \dots) / q_{ij} \\
 &= p_{ij} [1/(1 - q_{ij}) - 1 + \ln(1 - q_{ij})] / q_{ij} \\
 &= 1 + p_{ij} \ln(p_{ij}) / (1 - p_{ij})
 \end{aligned} \tag{A3}$$

All three expected values involve solutions of infinite series. The series solutions are given below.

One of the most celebrated series in mathematics is the *binomial series* (Salas and Hille 1974, pp. 533-534):

$$\sum_{k=0}^{\infty} (\pm 1)^k \binom{\alpha}{k} x^k = (1 \pm x)^\alpha, \quad x^2 < 1 \quad (\text{A4})$$

where $\binom{\alpha}{k}$ is the k th binomial coefficient. The expansion of the summation in (A4) looks like

$$1 \pm \alpha x + \frac{\alpha(\alpha-1)x^2}{2!} \pm \frac{\alpha(\alpha-1)(\alpha-2)x^3}{3!} + \dots \quad (\text{A5})$$

For the binomial $(1-x)^{-2}$ the following sequence is obtained

$$1 + 2x + 3x^2 + 4x^3 + \dots \quad (\text{A6})$$

which is the same sequence in Equation (A1).

To solve the power series embedded in Equation (A2) begin by noting that $\frac{d}{dx} -\ln u = -\frac{1}{u} \frac{du}{dx}$. We can sum the power series

$$\sum_{k=1}^{\infty} \frac{x^k}{k} = x + \frac{1}{2}x^2 + \frac{1}{3}x^3 + \frac{1}{4}x^4 + \dots, \quad x^2 < 1 \quad (\text{A7})$$

by setting

$$f(x) = \sum_{k=1}^{\infty} \frac{x^k}{k}, \quad x^2 < 1 \quad (\text{A8})$$

and applying the Differentiability of Power-Series Theorem (see Salas and Hille 1974, p. 526) to obtain

$$f'(x) = \sum_{k=1}^{\infty} \frac{kx^{k-1}}{k} = \sum_{k=1}^{\infty} x^{k-1} = \sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \quad (\text{A9})$$

the last sum being the well-known geometric series. Since $f(0) = 0$ and $f'(x) = \frac{1}{1-x}$ we must have

$$f(x) = -\ln(1-x) \quad (\text{A10})$$

To solve the power series embedded in Equation (A3) we begin by noting the summation formula:

$$\sum_{k=1}^{\infty} \frac{k-1}{k} x^k = \frac{1}{2}x^2 + \frac{2}{3}x^3 + \frac{3}{4}x^4 + \dots \quad (\text{A11})$$

We can take (A11) and rearrange it into two series with known values, and thereby arrive at the solution.

$$\begin{aligned} \sum_{k=1}^{\infty} (k-1) \frac{x^k}{k} &= \sum_{k=1}^{\infty} x^k - \sum_{k=1}^{\infty} \frac{x^k}{k} \\ &= \frac{1}{1-x} - 1 + \ln(1-x) \end{aligned} \quad (\text{A12})$$

Bounds for Γ_1

The index Γ_1 is bounded between 0 and 1. To show this I will make use of L'Hôpital's rule. L'Hôpital's rule says if $f(x)/g(x)$ has the indeterminate form $0/0$ or ∞/∞ at $x = c$, then

$$\lim_{x \rightarrow c} \frac{f(x)}{g(x)} = \lim_{x \rightarrow c} \frac{f'(x)}{g'(x)}, \quad g'(x) \neq 0. \quad \text{If } n = 1 \text{ then } p_{11} = 1 \text{ and } \Gamma_1 \text{ has the indeterminate form } 0/0,$$

so by L'Hôpital's rule we have:

$$\lim_{n \rightarrow 1} \sum_{i=1}^n \sum_{j=1}^n \frac{p_{ij}^2 \ln(p_{ij})}{p_{ij} - 1} = \lim_{p_{11} \rightarrow 1} \frac{p_{11} + 2p_{11} \ln(p_{11})}{1} = 1 \quad (\text{A13})$$

To establish the lower bound, consider that as $n \rightarrow \infty$ the p_{ij} 's $\rightarrow 0$. However, this results in the indeterminate form $0 \cdot \infty$. By writing $p_{ij}^2 \ln(p_{ij})$ as $\frac{\ln(p_{ij})}{1/p_{ij}^2}$ we obtain the indeterminate form ∞/∞ to which we can apply L'Hôpital's rule. Hence

$$\begin{aligned}
 \lim_{n \rightarrow \infty} \sum_{i=1}^n \sum_{j=1}^n \frac{p_{ij}^2 \ln(p_{ij})}{p_{ij}^2 - 1} &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \lim_{p_{ij} \rightarrow 0^+} \frac{\ln(p_{ij})}{1/p_{ij}^2} / (p_{ij}^2 - 1) \\
 &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \lim_{p_{ij} \rightarrow 0^+} \frac{1/p_{ij}}{-2/p_{ij}^3} / (p_{ij}^2 - 1) \\
 &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \lim_{p_{ij} \rightarrow 0^+} -\frac{p_{ij}^2}{2p_{ij}^2 - 2} \\
 &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{0}{2} = 0
 \end{aligned} \tag{A14}$$

APPENDIX B
FORTRAN CODE, EXAMPLE INPUT FILE, AND EXAMPLE OUTPUT
FOR PROGRAM CONTAGION

```
C *****  
C *  
C *  
C *  
C * THIS PROGRAM COMPUTES VALUES OF THE NEW GAMMA CONTAGION  
C * INDICES, ALONG WITH THE LI AND REYNOLDS INDEX RC2 AND  
C * RELATIVE EVENNESS. PORTIONS OF THIS PROGRAM WERE ADAPTED  
C * FROM HARBIN LI'S LSPA PROGRAM. INPUT IS A MATRIX OF INTEGERS  
C * WIH VALUES FROM ZERO TO ONE LESS THAN THE NUMBER OF PATCH  
C * TYPES. ROOK'S RULE IS USED FOR COMPUTING ADJACENCIES. THE  
C * PROGRAM PROMPTS FOR AN OUTPUT FILE NAME WHICH IT WRITES TO  
C * AND AN INPUT FILE NAME FROM WHICH IT READS. THE FIRST LINE  
C * MUST HAVE 3 VALUES SEPARATED BY BLANKS: THE NUMBER OF ROWS  
C * IN THE DATA MATRIX (NROW), THE NUMBER OF COLUMNS (NCOL), AND  
C * THE NUMBER OF LAND COVER CATEGORIES OR PATCH TYPES (NPATCH).  
C * THE NEXT LINE IS FOR A TITLE TO BE PRINTED IN THE OUTPUT,  
C * TO IDENTIFY IT. FOLLOWING THE TITLE LINE THE DATA MATRIX  
C * SHOULD BE LISTED WITH EACH VALUE SEPARATED BY A BLANK SPACE.  
C *  
C *  
C *  
C * PROGRAMMED BY BERNARD R. PARRESOL  
C * SCHOOL OF FORESTRY, WILDLIFE AND FISHERIES  
C * LOUISIANA STATE UNIVERSITY  
C * BATON ROUGE, LOUISIANA  
C *  
C *  
C * FEBRUARY, 1998  
C *  
C *  
C * ***** INPUT VARIABLES *****  
C * NROW = NUMBER OF ROWS IN LANDSCAPE MATRIX  
C * NCOL = NUMBER OF COLUMNS IN LANDSCAPE MATRIX  
C * NPATCH = NUMBER OF PATCH TYPES  
C * LMATRIX= LANDSCAPE MATRIX CONTAINING INTEGER VALUES REPRESENTING DIFFERENT PATCH TYPES, STARTING WITH 0  
C *  
C *  
C * ***** PROGRAM VARIABLES *****  
C * NRUN = INDEX VARIABLE  
C * LMWORK = WORKING COPY OF LMATRIX WITH BORDER CELLS OF -9  
C * NJC = MATRIX HOLDING Nij COUNTS  
C * NJT = VECTOR HOLDING Ni COUNTS  
C * P = VECTOR HOLDING PATCH TYPE PROPORTIONS (Pi)  
C * PROB = MATRIX HOLDING Pij VALUES FOR COMPUTING GAMMA & RC2  
C * GAMMA1 = NEW LANDSCAPE CONTAGION INDEX  
C * VGAM1 = VARIANCE OF GAMMA1  
C * GAMMA2 = ENTROPY LANDSCAPE CONTAGION INDEX  
C * VGAM2 = VARIANCE OF GAMMA2  
C * RC2 = LI AND REYNOLDS INDEX  
C * VRC2 = VARIANCE OF RC2  
C * RE = RELATIVE EVENNESS  
C *  
C *****  
C *  
C *  
C-----MAXIMUM PATCH TYPES CURRENTLY SET TO 10  
C-----MAXIMUM LANDSCAPE MATRIX CURRENTLY SET TO 60 BY 60  
C  
REAL*8 GAMMA1,GAMMA2,P(0:9),PROB(0:9,0:9),Pji,S0,SUM,T,RC2,RE,  
: VGAM1,VGAM2,VRC2  
INTEGER NJC(0:9,0:9),NJT(0:9)
```

```

      CHARACTER IFILE*24, OFILE*24, TITLE*70
      COMMON /INDEXES/ NROW, NCOL, NRUN
      COMMON /DATA/ LMATRIX(60,60)
      COMMON /WORK/ LMWORK(62,62)
C
      WRITE(*, '(//10X,A)') 'LANDSCAPE CONTAGION COMPUTATION'
      WRITE(*, '(/5X,A\')') 'ENTER OUTPUT FILE NAME: '
      READ(*, '(A)') OFILE
      OPEN(7, FILE=OFILE, STATUS='NEW')
      WRITE(*, '(/5X,A\')') 'INPUT DATA FILE NAME: '
      READ(*, '(A)') IFILE
      OPEN(5, FILE=IFILE)
C
C-----READ THE LANDSCAPE MATRIX VALUES
C
      READ(5, *) NROW, NCOL, NPATCH
      READ(5, '(A)') TITLE
      DO I=1, NROW
        READ(5, *) (LMATRIX(I, J), J=1, NCOL)
      END DO
      CLOSE(5, STATUS='KEEP')
C
C-----WRITE HEADER INFO INTO OUTPUT FILE
C
      WRITE(7, 10) TITLE
      10 FORMAT(20X, 'LANDSCAPE CONTAGION COMPUTATION'//, A/)
C
C-----CREATE NEW INDEX VARIABLE
C
      NRUN = NPATCH - 1
C
C-----COMPUTE PROPORTION OF EACH PATCH TYPE
C
      CALL PATCHPRP (P)
C
C-----CREATE WORKING MAP WITH ARTIFICIAL BORDER
C
      CALL BORDER
C
C-----COMPUTE ADJACENCIES
C
      CALL JOINCNT (NJC, NJT)
C
C-----COMPUTE THE Pij'S
C
      DO 20 I=0, NRUN
        DO 20 J=0, NRUN
          Pji = DFLOAT(NJC(I, J)) / DFLOAT(NJT(I))
          PROB(I, J) = P(I) * Pji
          IF (PROB(I, J) .LT. 1.D-9) PROB(I, J) = 1.D-9
        20 CONTINUE
C
C-----COMPUTE GAMMA1, VGAM1, GAMMA2, VGAM2, RC2, VRC2, AND RE
C
      GAMMA1 = 0.0D0
      VGAM1 = 0.0D0
      GAMMA2 = 0.0D0
      VGAM2 = 0.0D0

```

```

S0      = 0.0D0
RC2     = 0.0D0
VRC2    = 0.0D0
RE      = 0.0D0
DO 30 I=0, NRUN
  RE = RE + P(I) * P(I)
  DO 30 J=0, NRUN
    GAMMA1 = GAMMA1 + (PROB(I,J)*PROB(I,J)*DLOG(PROB(I,J))) /
:      (PROB(I,J) - 1.0D0)
    VGAM1  = VGAM1 + PROB(I,J)*(PROB(I,J)**2*(1+DLOG(PROB(I,J)))
: - PROB(I,J)*(1+2*DLOG(PROB(I,J))))**2 / (PROB(I,J)-1)**4
    S0     = S0 + (PROB(I,J)**3*(1+DLOG(PROB(I,J))-PROB(I,J)**2
:      * (1+2*DLOG(PROB(I,J)))) / (PROB(I,J)-1)**2
    GAMMA2 = GAMMA2 + PROB(I,J)*DLOG(PROB(I,J))
    VGAM2  = VGAM2 + PROB(I,J)*DLOG(PROB(I,J))*DLOG(PROB(I,J))
30 CONTINUE
C
  T = DFLOAT(NROW * NCOL)
  VGAM1 = (VGAM1 - S0*S0) / T
  VGAM2 = (VGAM2 - GAMMA2*GAMMA2) / T
  RC2   = 1.0D0 + GAMMA2 / (2.0D0 * DLOG(DFLOAT(NPATCH)))
  VRC2  = VGAM2 / (4.0D0 * DLOG(DFLOAT(NPATCH))**2)
  RE    = -DLOG(RE) / DLOG(DFLOAT(NPATCH))
C
  WRITE(7,40)
40 FORMAT(2X,'Ni j = '/')
C
  DO I=0, NRUN
    WRITE(7,50) (NJC(I,J),J=0,NRUN)
50  FORMAT(2X,10(I4,1X))
  END DO
C
  WRITE(7,60) (NJT(I),I=0,NRUN)
60 FORMAT(/2X,'Ni = ',10(I4,1X))
C
  WRITE(7,70) (P(I),I=0,NRUN)
70 FORMAT(/2X,5HPi = ,10(F6.4,1X))
C
  WRITE(7,80)
80 FORMAT(/2X,'Pi j = '/')
C
  DO I=0, NRUN
    WRITE(7,90) (PROB(I,J),J=0,NRUN)
90  FORMAT(2X,10(F6.4,1X))
  END DO
C
  DO 100 I=0, NRUN
    DO 100 J=0, NRUN
      SUM = SUM + PROB(I,J)
100 CONTINUE
C
  WRITE(7,110) SUM,RE,GAMMA1,VGAM1,GAMMA2,VGAM2,RC2,VRC2
110 FORMAT(/2X,'ΣPij = ',F6.4,' RE = ',F6.4,
: //2X,' Γ1 = ',F7.4,' V(Γ1) = ',F11.9,
: //2X,' Γ2 = ',F7.4,' V(Γ2) = ',F11.9,
: //2X,' RC2 = ',F7.4,' V(RC2) = ',F11.9)
C
  CLOSE(7,STATUS='KEEP')

```

```

      STOP '      PROGRAM CONTAGION TERMINATED'
      END

C
C
      SUBROUTINE PATCHPRP (P)
C
C
C *****
C *
C * SUBROUTINE PATCHPRP COMPUTES THE PROPORTION OF EACH PATCH TYPE *
C * IN THE LANDSCAPE MATRIX AS A SIMPLE RATIO OF TOTAL COUNT OF *
C * PATCH TYPE I PIXELS OVER TOTAL NUMBER OF PIXELS. *
C *
C *****
C
C
      REAL*8 P(0:9)
      INTEGER ICOUNT(0:9)
      COMMON /INDEXES/ NROW,NCOL,NRUN
      COMMON /DATA/ LMATRIX(60,60)
C
C-----NTC IS THE TOTAL NUMBER OF LATTICE POINTS OR PIXELS
C
      NTC = NROW * NCOL
C
C-----INITIALIZE COUNTER
C
      DO 10 I = 0, NRUN
          ICOUNT(I) = 0
      10 CONTINUE
C
C-----COUNT OCCURRENCE OF EACH PATCH TYPE
C
      DO 20 I=1, NROW
          DO 20 J=1, NCOL
              DO K = 0, NRUN
                  IF (LMATRIX(I,J) .EQ. K) ICOUNT(K)=ICOUNT(K)+1
              END DO
          20 CONTINUE
C
C-----COMPUTE PROPORTION OF EACH PATCH TYPE
C
      DO I = 0, NRUN
          P(I) = DFLOAT(ICOUNT(I)) / DFLOAT(NTC)
      END DO
      RETURN
      END
C
C
      SUBROUTINE BORDER
C
C
C *****
C *
C * SUBROUTINE BORDER CREATES A WORKING MATRIX AND SETS A BOUNDARY *
C * OF CELLS AROUND THE MATRIX TO -9. *
C *
C *****
C

```

```

C
C
      COMMON /INDEXES/ NROW,NCOL,NRUN
      COMMON /DATA/  LMATRIX(60,60)
      COMMON /WORK/  LMWORK(62,62)
C
C-----CREATE WORKING MAP
C
      DO 10 I=1, NROW+2
        DO 10 J=1, NCOL+2
          LMWORK(I,J) = -9
      10 CONTINUE
C
C-----COPY LANDSCAPE MATRIX TO WORKING MATRIX
C
      DO 20 I=2, NROW+1
        DO 20 J=2, NCOL+1
          LMWORK(I,J) = LMATRIX(I-1,J-1)
      20 CONTINUE
      RETURN
      END
C
C
      SUBROUTINE JOINCNT (NJC,NJT)
C
C
C *****
C *
C *  SUBROUTINE JOINCNT DETERMINES THE JOINCOUNT OF ALL THE Ni j
C *  ADJACENCIES AND THE Ni ADJACENCIES USING ROOK'S RULE.
C *
C *****
C
      INTEGER NJC(0:9,0:9),NJT(0:9)
      COMMON /INDEXES/ NROW,NCOL,NRUN
      COMMON /WORK/  LMWORK(62,62)
C
C-----INITIALIZE COUNTERS
C
      DO 10 I=0, NRUN
        DO 10 J=0, NRUN
          NJC(I,J) = 0
      10 CONTINUE
C
      DO 30 ID=0, NRUN
        DO 30 JD=0, NRUN
C
          DO 20 I=2, NROW+1
            DO 20 J=2, NCOL+1
              IF (LMWORK(I,J) .EQ. ID) THEN
                IF (LMWORK(I-1,J) .EQ. JD)  NJC(ID,JD) = NJC(ID,JD)+1
                IF (LMWORK(I,J+1) .EQ. JD)  NJC(ID,JD) = NJC(ID,JD)+1
                IF (LMWORK(I+1,J) .EQ. JD)  NJC(ID,JD) = NJC(ID,JD)+1
                IF (LMWORK(I,J-1) .EQ. JD)  NJC(ID,JD) = NJC(ID,JD)+1
              END IF
            20 CONTINUE
          30 CONTINUE

```

```
C
C-----CORRECT DOUBLE COUNT IN THE DIAGONAL OF MATRIX NJC
C
      DO I=0, NRUN
        NJC(I,I) = NJC(I,I) / 2
      END DO
C
C-----INITIALIZE COUNTERS
C
      DO 40 I=0, NRUN
        NJT(I) = 0
      40 CONTINUE
C
C-----COMPUTE TOTAL # OF ADJACENCIES BETWEEN TYPE i AND ALL TYPES
C
      DO 50 I=0, NRUN
        DO 50 J=0, NRUN
          NJT(I) = NJT(I) + NJC(I,J)
        50 CONTINUE
      RETURN
      END
```


Example Input File

```

4 5 3
Test of Program CONTAGION
2 0 1 1 2
1 0 0 1 0
0 1 2 0 1
1 1 2 1 2

```

Example Output

```

                                LANDSCAPE CONTAGION COMPUTATION

Test of Program CONTAGION

Nij =

      2    13    4
    13     4    7
      4     7    1

Ni =   19   24   12

Pi =   .3000  .4500  .2500

Pij =

    .0316  .2053  .0632
    .2438  .0750  .1313
    .0833  .1458  .0208

ΣΣPij = 1.0000 RE = .9427

Γ1 =   .3347  V(Γ1) = .001360119

Γ2 =  -1.9820  V(Γ2) = .017142236

RC2 =   .0980  V(RC2) = .003550737

```

APPENDIX C

**FORTRAN CODE, EXAMPLE INPUT FILE, AND EXAMPLE OUTPUT
FOR PROGRAM VAREST**

```

C *****
C *
C *
C *          PROGRAM VAREST
C *          (nonparametric VARIance ESTimation)
C *
C *
C *      REPLICATE SUBSERIES ARE USED FOR NONPARAMETRIC VARIANCE
C *      ESTIMATION OF THE NEW GAMMA LANDSCAPE CONTAGION INDICES.
C *      SEE SHERMAN, M. AND CARLSTEIN, E. 1994. NONPARAMETRIC
C *      ESTIMATION OF THE MOMENTS OF A GENERAL STATISTIC COMPUTED
C *      FROM SPATIAL DATA. JOURNAL OF THE AMERICAN STATISTICAL
C *      ASSOCIATION 89: 496-500.
C *
C *
C *      THIS PROGRAM CREATES NONOVERLAPPING SUBRECTANGLES OF SIZE L1
C *      BY L2 FROM THE INPUT LANDSCAPE MATRIX AND COMPUTES VALUES OF
C *      THE NEW GAMMA CONTAGION INDICES AND RELATIVE EVENNESS ON ALL
C *      THE SUBSETS OF THE DATA. PORTIONS OF THIS PROGRAM WERE
C *      ADAPTED FROM HARBIN LI'S LSPA PROGRAM. INPUT IS A MATRIX OF
C *      INTEGERS WITH VALUES FROM ZERO TO ONE LESS THAN THE NUMBER OF
C *      PATCH TYPES. ROOK'S RULE IS USED FOR COMPUTING ADJACENCIES.
C *      THE PROGRAM PROMPTS FOR AN OUTPUT FILE NAME WHICH IT WRITES
C *      TO AND AN INPUT FILE NAME FROM WHICH IT READS. THE FIRST
C *      LINE MUST HAVE 5 VALUES SEPARATED BY BLANKS: THE NUMBER OF
C *      ROWS IN THE DATA MATRIX (NROW), THE NUMBER OF COLUMNS
C *      (NCOL), THE NUMBER OF LAND COVER CATEGORIES OR PATCH TYPES
C *      (NPATCH), THE LENGTH OF SUBRECTANGLE SIDE 1 (L1), AND LENGTH
C *      OF SIDE 2 (L2). NOTE THAT L1 MUST BE A MULTIPLE OF NROW AND
C *      L2 MUST BE A MULTIPLE OF NCOL. THE NEXT LINE IS FOR A TITLE
C *      TO BE PRINTED IN THE OUTPUT, TO IDENTIFY IT. FOLLOWING THE
C *      TITLE LINE THE DATA MATRIX SHOULD BE LISTED WITH EACH VALUE
C *      SEPARATED BY A BLANK SPACE.
C *
C *
C *
C *          PROGRAMMED BY BERNARD R. PARRESOL
C *          SCHOOL OF FORESTRY, WILDLIFE AND FISHERIES
C *          LOUISIANA STATE UNIVERSITY
C *          BATON ROUGE, LOUISIANA
C *
C *
C *          FEBRUARY, 1998
C *
C *
C *          ***** INPUT VARIABLES *****
C *      NROW   =  NUMBER OF ROWS IN LANDSCAPE MATRIX
C *      NCOL   =  NUMBER OF COLUMNS IN LANDSCAPE MATRIX
C *      NPATCH =  NUMBER OF PATCH TYPES
C *      L1     =  LENGTH OF SIDE 1 OF EACH SUBRECTANGLE
C *      L2     =  LENGTH OF SIDE 2 OF EACH SUBRECTANGLE
C *      LMATRIX=  LANDSCAPE MATRIX CONTAINING INTEGER VALUES REPRESENTING
C *                DIFFERENT PATCH TYPES, STARTING WITH 0
C *      TITLE  =  OUTPUT TITLE OF UP TO 70 CHARACTERS (CAN BE BLANKS)
C *
C *
C *          ***** PROGRAM VARIABLES *****
C *      NRUN   =  INDEX VARIABLE FOR PATCH TYPES
C *      LMSUB  =  MATRIX HOLDING 1th SUBSQUARE OF LMATRIX
C *      LMWORK =  WORKING COPY OF LMSUB WITH BORDER CELLS OF -9
C *      NJC    =  MATRIX HOLDING Nij COUNTS
C *      NJT    =  VECTOR HOLDING Ni COUNTS
C *      P      =  VECTOR HOLDING PATCH TYPE PROPORTIONS (Pi)
C *      PROB   =  MATRIX HOLDING Pij VALUES FOR COMPUTING GAMMA & RC2
C *      GAMMA1 =  NEW LANDSCAPE CONTAGION INDEX
C *      GAMMA2 =  ENTROPY LANDSCAPE CONTAGION INDEX

```

```

C      *      RE      =  RELATIVE EVENNESS      *
C      *      SG11    =  SUMMATION FOR 1st MOMENT OF GAMMA1      *
C      *      SG12    =  SUMMATION FOR 2nd MOMENT OF GAMMA1      *
C      *      SG21    =  SUMMATION FOR 1st MOMENT OF GAMMA2      *
C      *      SG22    =  SUMMATION FOR 2nd MOMENT OF GAMMA2      *
C      *      VG1     =  VARIANCE OF GAMMA1      *
C      *      VG2     =  VARIANCE OF GAMMA2      *
C      *
C      *****
C
C
C
C-----MAXIMUM PATCH TYPES SET TO 10
C-----MAXIMUM LANDSCAPE MATRIX SET TO 60 BY 60
C
      REAL*8 GAMMA1,GAMMA2,P(0:9),PROB(0:9,0:9),Pji,RE,SG11,SG12,SG21,
      :      SG22,VG1,VG2
      INTEGER LMATRIX(60,60),NJC(0:9,0:9),NJT(0:9)
      CHARACTER IFILE*24,OFILE*24,TITLE*70
      COMMON /INDEXES/ L1,L2,NRUN
      COMMON /DATA/ LMSUB(60,60)
      COMMON /WORK/ LMWORK(62,62)
C
      WRITE(*,'(//10X,A)') 'Nonparametric Variance Estimation of Landsc
:ape Contagion'
      WRITE(*,'(/5X,A)') 'ENTER OUTPUT FILE NAME: '
      READ(*,'(A)') OFILE
      OPEN(7,FILE=OFILE,STATUS='NEW')
      WRITE(*,'(/5X,A)') 'INPUT DATA FILE NAME: '
      READ(*,'(A)') IFILE
      OPEN(5,FILE=IFILE)
C
C-----READ THE LANDSCAPE MATRIX VALUES
C
      READ(5,*) NROW,NCOL,NPATCH,L1,L2
      READ(5,'(A)') TITLE
      DO I=1,NROW
          READ(5,*)(LMATRIX(I,J),J=1,NCOL)
      END DO
      CLOSE(5,STATUS='KEEP')
C
C-----WRITE HEADER INFO INTO OUTPUT FILE
C
      WRITE(7,10) TITLE
      10 FORMAT(7X,'Nonparametric Variance Estimation of Landscape Contagio
:n'//,A//,'SUBREGION',4X,'Γ1',6X,'Γ2',6X,'RE')
C
      WRITE(*,'(/5X,A\,/)') 'PROGRAM WORKING . . .'
C
C-----CREATE NEW INDEX VARIABLE
C
      NRUN = NPATCH - 1
C
C-----INITIALIZE CONTROL, LOOPING, AND SUMMING VARIABLES
C
      KOUNTER = 0
      MAX = (NROW * NCOL) / (L1*L2)
      MACROSS = NCOL / L2
      NC = 1
      NR = 1

```

```

      SG11 = 0.0D0
      SG12 = 0.0D0
      SG21 = 0.0D0
      SG22 = 0.0D0
C
C-----START SUBSERIES ROUTINE
C
      DO 60 ITER=1, MAX
C
C-----CREATE ith NONOVERLAPPING SUBRECTANGLE
C
      KOUNTER = KOUNTER + 1
C
      IF (KOUNTER .GT. MACROSS) THEN
        KOUNTER = 1
        NR = NR + L1
        NC = 1
      END IF
C
      ID = 0
      DO 20 I=NR, NR+L1-1
        ID = ID + 1
        JD = 0
        DO 20 J=NC, NC+L2-1
          JD = JD + 1
          LMSUB(ID,JD) = LMATRIX(I,J)
20    CONTINUE
      NC = NC + L2
C
C-----COMPUTE PROPORTION OF EACH PATCH TYPE
C
      CALL PATCHPRP (P)
C
C-----CREATE WORKING MAP WITH ARTIFICAL BORDER
C
      CALL BORDER
C
C-----COMPUTE ADJACENCIES
C
      CALL JOINCNT (NJC,NJT)
C
C-----COMPUTE THE Pij'S
C
      IADJUST = 0
      DO 30 I=0, NRUN
        IF (NJT(I) .EQ. 0) THEN
          IADJUST = IADJUST + 1
          NJT(I) = 1
        END IF
        DO 30 J=0, NRUN
          Pji = DFLOAT(NJC(I,J)) / DFLOAT(NJT(I))
          PROB(I,J) = P(I) * Pji
          IF (PROB(I,J) .LT. 1.D-9) PROB(I,J) = 1.D-9
30    CONTINUE
C
C-----COMPUTE GAMMA1 AND GAMMA2 AND RE
C
      IF ((NPATCH-IADJUST) .EQ. 1) THEN
        GAMMA1 = 1.0D0

```

```

        GAMMA2 = 0.0D0
        RE      = 1.0D0
    ELSE
        GAMMA1 = 0.0D0
        GAMMA2 = 0.0D0
        RE      = 0.0D0
        DO 40 I=0, NRUN
            RE = RE + P(I) * P(I)
            DO 40 J=0, NRUN
                GAMMA1 = GAMMA1 + (PROB(I,J)*PROB(I,J)*DLOG(PROB(I,J))) /
:                (PROB(I,J) - 1.0D0)
                GAMMA2 = GAMMA2 + PROB(I,J)*DLOG(PROB(I,J))
40      CONTINUE
C
        RE = -DLOG(RE) / DLOG(DFLOAT(NPATCH-IADJUST))
    END IF
C
C-----SUMMATIONS FOR 1st AND 2nd NONCENTRAL MOMENT ESTIMATION
C
        SG11 = SG11 + GAMMA1
        SG21 = SG21 + GAMMA2
        SG12 = SG12 + GAMMA1*GAMMA1
        SG22 = SG22 + GAMMA2*GAMMA2
C
        WRITE(7,50) ITER,GAMMA1,GAMMA2,RE
50      FORMAT(2X,I4,4X,F6.4,2X,F7.4,1X,F6.4)
C
60 CONTINUE
C
C-----NONPARAMETRIC VARIANCE ESTIMATES
C
        VG1 = SG12/DFLOAT(MAX) - (SG11/DFLOAT(MAX))**2
        VG2 = SG22/DFLOAT(MAX) - (SG21/DFLOAT(MAX))**2
C
        WRITE(7,70) VG1,VG2
70      FORMAT(/2X,'V(Γ1) =',F10.8,2X,'V(Γ2) =',F10.8)
C
        CLOSE(7,STATUS='KEEP')
        STOP '      PROGRAM VAREST TERMINATED'
        END
C
C
C      SUBROUTINE PATCHPRP (P)
C
C
C      *****
C      *
C      * SUBROUTINE PATCHPRP COMPUTES THE PROPORTION OF EACH PATCH TYPE *
C      * IN THE LANDSCAPE MATRIX AS A SIMPLE RATIO OF TOTAL COUNT OF *
C      * PATCH TYPE i PIXELS OVER TOTAL NUMBER OF PIXELS. *
C      *
C      *****
C
C
C      REAL*8 P(0:9)
C      INTEGER ICOUNT(0:9)
C      COMMON /INDEXES/ L1,L2,NRUN
C      COMMON /DATA/ LMSUB(60,60)
C

```

```

C-----NTC IS THE TOTAL NUMBER OF LATTICE POINTS OR PIXELS
C
      NTC = L1 * L2
C
C-----INITIALIZE COUNTER
C
      DO 10 I = 0, NRUN
        ICOUNT(I) = 0
      10 CONTINUE
C
C-----COUNT OCCURRENCE OF EACH PATCH TYPE
C
      DO 20 I=1, L1
        DO 20 J=1, L2
          DO K = 0, NRUN
            IF (LMSUB(I,J) .EQ. K) ICOUNT(K)=ICOUNT(K)+1
          END DO
        20 CONTINUE
C
C-----COMPUTE PROPORTION OF EACH PATCH TYPE
C
      DO I = 0, NRUN
        P(I) = DFLOAT(ICOUNT(I)) / DFLOAT(NTC)
      END DO
      RETURN
      END
C
C
      SUBROUTINE BORDER
C
C
C *****
C *
C * SUBROUTINE BORDER CREATES A WORKING MATRIX AND SETS A BOUNDARY *
C * OF CELLS AROUND THE MATRIX TO -9. *
C *
C *****
C
      COMMON /INDEXES/ L1,L2,NRUN
      COMMON /DATA/ LMSUB(60,60)
      COMMON /WORK/ LMWORK(62,62)
C
C-----CREATE WORKING MAP
C
      DO 10 I=1, L1+2
        DO 10 J=1, L2+2
          LMWORK(I,J) = -9
        10 CONTINUE
C
C-----COPY LANDSCAPE MATRIX TO WORKING MATRIX
C
      DO 20 I=2, L1+1
        DO 20 J=2, L2+1
          LMWORK(I,J) = LMSUB(I-1,J-1)
        20 CONTINUE
      RETURN
      END
C

```

```

C
      SUBROUTINE JOINCNT (NJC,NJT)
C
C
C *****
C *
C *   SUBROUTINE JOINCNT DETERMINES THE JOINCOUNT OF ALL THE Nij
C *   ADJACENCIES AND THE Ni ADJACENCIES USING ROOK'S RULE.
C *
C *****
C
C
      INTEGER NJC(0:9,0:9),NJT(0:9)
      COMMON /INDEXES/ L1,L2,NRUN
      COMMON /WORK/ LMWOPK(62,62)
C
C-----INITIALIZE COUNTERS
C
      DO 10 I=0, NRUN
        DO 10 J=0, NRUN
          NJC(I,J) = 0
        10 CONTINUE
C
      DO 30 ID=0, NRUN
        DO 30 JD=0, NRUN
C
          DO 20 I=2, L1+1
            DO 20 J=2, L2+1
              IF (LMWORK(I,J) .EQ. ID) THEN
                IF (LMWORK(I-1,J) .EQ. JD) NJC(ID,JD) = NJC(ID,JD)+1
                IF (LMWORK(I,J+1) .EQ. JD) NJC(ID,JD) = NJC(ID,JD)+1
                IF (LMWORK(I+1,J) .EQ. JD) NJC(ID,JD) = NJC(ID,JD)+1
                IF (LMWORK(I,J-1) .EQ. JD) NJC(ID,JD) = NJC(ID,JD)+1
              END IF
            20 CONTINUE
          30 CONTINUE
C
C-----CORRECT DOUBLE COUNT IN THE DIAGONAL OF MATRIX NJC
C
      DO I=0, NRUN
        NJC(I,I) = NJC(I,I) / 2
      END DO
C
C-----INITIALIZE COUNTERS
C
      DO 40 I=0, NRUN
        NJT(I) = 0
      40 CONTINUE
C
C-----COMPUTE TOTAL # OF ADJACENCIES BETWEEN TYPE i AND ALL TYPES
C
      DO 50 I=0, NRUN
        DO 50 J=0, NRUN
          NJT(I) = NJT(I) + NJC(I,J)
        50 CONTINUE
      RETURN
      END

```


Example Input File

```

12 12 5 4 4
Test of Program VAREST
 3 3 2 4 3 2 2 4 3 3 0 2
 3 0 3 3 4 3 3 4 2 2 3 3
 3 0 4 4 1 3 2 4 3 4 0 3
 0 1 3 0 1 0 0 0 2 2 2 3
 4 3 1 3 1 0 4 0 3 0 3 3
 2 2 2 3 4 2 3 1 3 3 0 3
 4 1 2 3 3 0 2 4 2 2 4 2
 0 4 1 0 4 4 3 3 3 3 4 0
 2 3 4 0 2 0 0 2 3 2 2 4
 1 1 3 4 0 3 3 2 1 0 2 4
 3 4 4 2 3 3 3 1 2 3 1 0
 3 2 2 3 0 2 2 4 2 2 3 0

```

Example Output**Nonparametric Variance Estimation of Landscape Contagion**

Test of Program VAREST

SUBREGION	Γ_1	Γ_2	RE
1	.2246	-2.6096	.7546
2	.1676	-2.9990	.9669
3	.2708	-2.3467	.7541
4	.1716	-2.9187	.9669
5	.1839	-2.8606	.9225
6	.2621	-2.4021	.7869
7	.2082	-2.6923	.8811
8	.2206	-2.6186	.8237
9	.1818	-2.8528	.8811

$V(\Gamma_1) = .00127606$ $V(\Gamma_2) = .04608492$

APPENDIX D

**FORTRAN CODE, EXAMPLE INPUT FILE, AND EXAMPLE OUTPUT
FOR PROGRAM REPHIST**

```

C *****
C *
C *
C *          PROGRAM REPHIST
C *          (REPLICATE HISTOGRAM)
C *
C *
C *  A "REPLICATE HISTOGRAM" IS A SIMPLE DIAGNOSTIC TOOL FOR
C *  DESCRIBING THE SAMPLING DISTRIBUTION OF A GENERAL STATISTIC.
C *  SEE SHERMAN, M. AND CARLSTEIN, E. 1996. REPLICATE HISTOGRAMS.
C *  JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION 91: 566-576.
C *
C *
C *  THIS PROGRAM CREATES ALL POSSIBLE SQUARE SUBSETS WITH LENGTH
C *  OF SIDE L FROM THE INPUT LANDSCAPE MATRIX AND COMPUTES VALUES
C *  OF THE NEW GAMMA CONTAGION INDICES AND RELATIVE EVENNESS ON
C *  ALL THE SUBSETS OF THE DATA. PORTIONS OF THIS PROGRAM WERE
C *  ADAPTED FROM HARBIN LI'S LSPA PROGRAM. INPUT IS A MATRIX OF
C *  INTEGERS WITH VALUES FROM ZERO TO ONE LESS THAN THE NUMBER OF
C *  PATCH TYPES. ROOK'S RULE IS USED FOR COMPUTING ADJACENCIES.
C *  THE PROGRAM PROMPTS FOR AN OUTPUT FILE NAME WHICH IT WRITES
C *  TO AND AN INPUT FILE NAME FROM WHICH IT READS. THE FIRST
C *  LINE MUST HAVE 4 VALUES SEPARATED BY BLANKS: THE NUMBER OF
C *  ROWS IN THE DATA MATRIX (NROW), THE NUMBER OF COLUMNS (NCOL),
C *  THE NUMBER OF LAND COVER CATEGORIES OR PATCH TYPES (NPATCH),
C *  AND THE BLOCK LENGTH L. THE NEXT LINE IS FOR A TITLE TO BE
C *  PRINTED IN THE OUTPUT, TO IDENTIFY IT. FOLLOWING THE TITLE
C *  LINE THE DATA MATRIX SHOULD BE LISTED WITH EACH VALUE SEPAR-
C *  ATED BY A BLANK SPACE. THE VALUES IN THE OUTPUT FILE ARE
C *  USED FOR PLOTTING KERNEL SMOOTHED REPLICATE HISTOGRAMS OF
C *  GAMMA1 AND GAMMA2.
C *
C *
C *
C *          PROGRAMMED BY BERNARD R. PARRESOL
C *          SCHOOL OF FORESTRY, WILDLIFE AND FISHERIES
C *          LOUISIANA STATE UNIVERSITY
C *          BATON ROUGE, LOUISIANA
C *
C *
C *          FEBRUARY, 1998
C *
C *
C *          ***** INPUT VARIABLES *****
C *
C *  NROW   =  NUMBER OF ROWS IN LANDSCAPE MATRIX
C *  NCOL   =  NUMBER OF COLUMNS IN LANDSCAPE MATRIX
C *  NPATCH =  NUMBER OF PATCH TYPES
C *  L      =  LENGTH OF SIDE OF EACH SUBSQUARE
C *  LMATRIX=  LANDSCAPE MATRIX CONTAINING INTEGER VALUES REPRESENTING
C *             DIFFERENT PATCH TYPES, STARTING WITH 0
C *  TITLE  =  OUTPUT TITLE OF UP TO 70 CHARACTERS (CAN BE BLANKS)
C *
C *
C *          ***** PROGRAM VARIABLES *****
C *
C *  NRUN   =  INDEX VARIABLE FOR PATCH TYPES
C *  LMSUB  =  MATRIX HOLDING ith SUBSQUARE OF LMATRIX
C *  LMWORK =  WORKING COPY OF LMSUB WITH BORDER CELLS OF -9
C *  NJC    =  MATRIX HOLDING Nij COUNTS
C *  NJT    =  VECTOR HOLDING Ni COUNTS
C *  P      =  VECTOR HOLDING PATCH TYPE PROPORTIONS (Pi)
C *  PROB   =  MATRIX HOLDING Pij VALUES FOR COMPUTING GAMMA & RC2
C *  GAMMA1 =  NEW LANDSCAPE CONTAGION INDEX
C *  GAMMA2 =  ENTROPY LANDSCAPE CONTAGION INDEX
C *  RE     =  RELATIVE EVENNESS
C *
C *****

```

```

C
C
C-----MAXIMUM PATCH TYPES SET TO 10
C-----MAXIMUM LANDSCAPE MATRIX SET TO 60 BY 60
C
      REAL*8 GAMMA1,GAMMA2,P(0:9),PROB(0:9,0:9),Pji,RE
      INTEGER LMATRIX(60,60),NJC(0:9,0:9),NJT(0:9)
      CHARACTER IFILE*24,OFILE*24,TITLE*70
      COMMON /INDEXES/ L,NRUN
      COMMON /DATA/ LMSUB(60,60)
      COMMON /WORK/ LMWORK(62,62)
C
      WRITE(*,'(//10X,A)') 'Replicate Histogram Analysis of Landscape Co
:ntagion'
      WRITE(*,'(/5X,A)') 'ENTER OUTPUT FILE NAME: '
      READ(*,'(A)') OFILE
      OPEN(7,FILE=OFILE,STATUS='NEW')
      WRITE(*,'(/5X,A)') 'INPUT DATA FILE NAME: '
      READ(*,'(A)') IFILE
      OPEN(5,FILE=IFILE)
C
C-----READ THE LANDSCAPE MATRIX VALUES
C
      READ(5,*) NROW,NCOL,NPATCH,L
      READ(5,'(A)') TITLE
      DO I=1,NROW
        READ(5,*)(LMATRIX(I,J),J=1,NCOL)
      END DO
      CLOSE(5,STATUS='KEEP')
C
C-----WRITE HEADER INFO INTO OUTPUT FILE
C
      WRITE(7,10) TITLE
      10 FORMAT(10X,'Replicate Histogram Analysis of Landscape Contagion'
:    //,A//,'SUBSQUARE',4X,'Γ1',6X,'Γ2',6X,'RE')
C
      WRITE(*,'(/5X,A\,/)') 'PROGRAM WORKING . . .'
C
C-----CREATE NEW INDEX VARIABLE
C
      NRUN = NPATCH - 1
C
C-----INITIALIZE CONTROL AND LOOPING VARIABLES
C
      KOUNTER = 0
      MAX = (NROW - L + 1) * (NCOL - L + 1)
      MACROSS = NCOL - L + 1
      NC = 1
      NR = 1
C
C-----START SUBSERIES ROUTINE
C
      DO 60 ITER=1, MAX
C
C-----CREATE ith SUBSQUARE
C
      KOUNTER = KOUNTER + 1
C

```

```

      IF (KOUNTER .GT. MACROSS) THEN
        KOUNTER = 1
        NR = NR + 1
        NC = 1
      END IF
C
      ID = 0
      DO 20 I=NR, NR+L-1
        ID = ID + 1
        JD = 0
        DO 20 J=NC, NC+L-1
          JD = JD + 1
          LMSUB(ID,JD) = LMATRIX(I,J)
20    CONTINUE
      NC = NC + 1
C
C-----COMPUTE PROPORTION OF EACH PATCH TYPE
C
      CALL PATCHPRP (P)
C
C-----CREATE WORKING MAP WITH ARTIFICIAL BORDER
C
      CALL BORDER
C
C-----COMPUTE ADJACENCIES
C
      CALL JOINCNT (NJC,NJT)
C
C-----COMPUTE THE Pij'S
C
      IADJUST = 0
      DO 30 I=0, NRUN
        IF (NJT(I) .EQ. 0) THEN
          IADJUST = IADJUST + 1
          NJT(I) = 1
        END IF
        DO 30 J=0, NRUN
          Pji = DFLOAT(NJC(I,J)) / DFLOAT(NJT(I))
          PROB(I,J) = P(I) * Pji
          IF (PROB(I,J) .LT. 1.D-9) PROB(I,J) = 1.D-9
30    CONTINUE
C
C-----COMPUTE GAMMA1 AND GAMMA2 AND RE
C
      IF ((NPATCH-IADJUST) .EQ. 1) THEN
        GAMMA1 = 1.0D0
        GAMMA2 = 0.0D0
        RE = 1.0D0
      ELSE
        GAMMA1 = 0.0D0
        GAMMA2 = 0.0D0
        RE = 0.0D0
        DO 40 I=0, NRUN
          RE = RE + P(I) * P(I)
          DO 40 J=0, NRUN
            GAMMA1 = GAMMA1 + (PROB(I,J)*PROB(I,J)*DLOG(PROB(I,J))) /
:              (PROB(I,J) - 1.0D0)
            GAMMA2 = GAMMA2 + PROB(I,J)*DLOG(PROB(I,J))
40    CONTINUE

```

```

C          RE = -DLOG(RE) / DLOG(DFLOAT(NPATCH-IADJUST))
C          END IF
C
C          WRITE(7,50) ITER,GAMMA1,GAMMA2,RE
50      FORMAT(2X,I4,4X,F6.4,2X,F7.4,1X,F6.4)
C
C      60 CONTINUE
C
C      CLOSE(7,STATUS='KEEP')
C      STOP '      PROGRAM REPHIST TERMINATED'
C      END
C
C
C      SUBROUTINE PATCHPRP (P)
C
C      *****
C      *
C      * SUBROUTINE PATCHPRP COMPUTES THE PROPORTION OF EACH PATCH TYPE *
C      * IN THE LANDSCAPE MATRIX AS A SIMPLE RATIO OF TOTAL COUNT OF *
C      * PATCH TYPE i PIXELS OVER TOTAL NUMBER OF PIXELS. *
C      *
C      *****
C
C      REAL*8 P(0:9)
C      INTEGER ICOUNT(0:9)
C      COMMON /INDEXES/ L,NRUN
C      COMMON /DATA/ LMSUB(60,60)
C
C      C-----NTC IS THE TOTAL NUMBER OF LATTICE POINTS OR PIXELS
C
C      NTC = L * L
C
C      C-----INITIALIZE COUNTER
C
C      DO 10 I = 0, NRUN
C          ICOUNT(I) = 0
C      10 CONTINUE
C
C      C-----COUNT OCCURRENCE OF EACH PATCH TYPE
C
C      DO 20 I=1, L
C          DO 20 J=1, L
C              DO K = 0, NRUN
C                  IF (LMSUB(I,J) .EQ. K) ICOUNT(K)=ICOUNT(K)+1
C              END DO
C      20 CONTINUE
C
C      C-----COMPUTE PROPORTION OF EACH PATCH TYPE
C
C      DO I = 0, NRUN
C          P(I) = DFLOAT(ICOUNT(I)) / DFLOAT(NTC)
C      END DO
C      RETURN
C      END
C
C

```

```

      SUBROUTINE BORDER
C
C
C *****
C *
C *   SUBROUTINE BORDER CREATES A WORKING MATRIX AND SETS A BOUNDARY *
C *   OF CELLS AROUND THE MATRIX TO -9. *
C *
C *****
C
C
C      COMMON /INDEXES/ L,NRUN
C      COMMON /DATA/ LMSUB(60,60)
C      COMMON /WORK/ LMWORK(62,62)
C
C-----CREATE WORKING MAP
C
C      DO 10 I=1, L+2
C        DO 10 J=1, L+2
C          LMWORK(I,J) = -9
C      10 CONTINUE
C
C-----COPY LANDSCAPE MATRIX TO WORKING MATRIX
C
C      DO 20 I=2, L+1
C        DO 20 J=2, L+1
C          LMWORK(I,J) = LMSUB(I-1,J-1)
C      20 CONTINUE
C      RETURN
C      END
C
C
C      SUBROUTINE JOINCNT (NJC,NJT)
C
C
C *****
C *
C *   SUBROUTINE JOINCNT DETERMINES THE JOINCOUNT OF ALL THE Ni,j *
C *   ADJACENCIES AND THE Ni ADJACENCIES USING ROOK'S RULE. *
C *
C *****
C
C
C      INTEGER NJC(0:9,0:9),NJT(0:9)
C      COMMON /INDEXES/ L,NRUN
C      COMMON /WORK/ LMWORK(62,62)
C
C-----INITIALIZE COUNTERS
C
C      DO 10 I=0, NRUN
C        DO 10 J=0, NRUN
C          NJC(I,J) = 0
C      10 CONTINUE
C
C
C      DO 30 ID=0, NRUN
C        DO 30 JD=0, NRUN
C
C          DO 20 I=2, L+1
C            DO 20 J=2, L+1

```

```

        IF (LMWORK(I,J) .EQ. ID) THEN
            IF (LMWORK(I-1,J) .EQ. JD) NJC(ID,JD) = NJC(ID,JD)+1
            IF (LMWORK(I,J+1) .EQ. JD) NJC(ID,JD) = NJC(ID,JD)+1
            IF (LMWORK(I+1,J) .EQ. JD) NJC(ID,JD) = NJC(ID,JD)+1
            IF (LMWORK(I,J-1) .EQ. JD) NJC(ID,JD) = NJC(ID,JD)+1
        END IF
    20 CONTINUE
    30 CONTINUE
C
C-----CORRECT DOUBLE COUNT IN THE DIAGONAL OF MATRIX NJC
C
    DO I=0, NRUN
        NJC(I,I) = NJC(I,I) / 2
    END DO
C
C-----INITIALIZE COUNTERS
C
    DO 40 I=0, NRUN
        NJT(I) = 0
    40 CONTINUE
C
C-----COMPUTE TOTAL # OF ADJACENCIES BETWEEN TYPE 1 AND ALL TYPES
C
    DO 50 I=0, NRUN
        DO 50 J=0, NRUN
            NJT(I) = NJT(I) + NJC(I,J)
        50 CONTINUE
    RETURN
END

```


Example Input File

```

12 12 5 9
Test of Program REPHIST
  3 3 2 4 3 2 2 4 3 3 0 2
  3 0 3 3 4 3 3 4 2 2 3 3
  3 0 4 4 1 3 2 4 3 4 0 3
  0 1 3 0 1 0 0 0 2 2 2 3
  4 3 1 3 1 0 4 0 3 0 3 3
  2 2 2 3 4 2 3 1 3 3 0 3
  4 1 2 3 3 0 2 4 2 2 4 2
  0 4 1 0 4 4 3 3 3 3 4 0
  2 3 4 0 2 0 0 2 3 2 2 4
  1 1 3 4 0 3 3 2 1 0 2 4
  3 4 4 2 3 3 3 1 2 3 1 0
  3 2 2 3 0 2 2 4 2 2 3 0

```

Example Output

Replicate Histogram Analysis of Landscape Contagion

Test of Program REPHIST

SUBSQUARE	Γ_1	Γ_2	RE
1	.1604	-3.0578	.9265
2	.1645	-3.0415	.9231
3	.1687	-3.0069	.9156
4	.1794	-2.9150	.8749
5	.1517	-3.1067	.9496
6	.1537	-3.0950	.9453
7	.1601	-3.0541	.9257
8	.1690	-2.9879	.8937
9	.1482	-3.1264	.9620
10	.1512	-3.1095	.9496
11	.1540	-3.0918	.9453
12	.1630	-3.0362	.9140
13	.1489	-3.1205	.9575
14	.1559	-3.0815	.9324
15	.1616	-3.0471	.9165
16	.1674	-3.0096	.8993

APPENDIX E
TABLES OF INDEX VALUES FROM SIMULATED LANDSCAPES

Table E1. Index values from 27 simulated landscapes under three spatial arrangements and increasing number of patch types/evenness gradient.

Spatial pattern	Number of patch types	RE	Γ_1	Γ_2	RC_2	D_2^*
Aggregated	2	0.2863	0.8255	-0.5339	0.6149	0.8524
Aggregated	3	0.3782	0.6649	-1.0459	0.5240	1.1513
Aggregated	4	0.4717	0.5315	-1.5163	0.4531	1.2563
Aggregated	5	0.5693	0.4337	-1.9075	0.4074	1.3114
Aggregated	6	0.6720	0.3613	-2.2132	0.3824	1.3703
Aggregated	7	0.7781	0.2903	-2.5273	0.3506	1.3645
Aggregated	8	0.8813	0.2149	-2.8925	0.3045	1.2664
Aggregated	9	0.9650	0.1730	-3.1592	0.2811	1.2353
Aggregated	10	1.0000	0.1548	-3.2927	0.2850	1.3125
Random	2	0.2863	0.7163	-0.7661	0.4474	0.6202
Random	3	0.3782	0.5196	-1.4339	0.3474	0.7633
Random	4	0.4717	0.3609	-2.0309	0.2675	0.7417
Random	5	0.5693	0.2591	-2.5819	0.1979	0.6370
Random	6	0.6720	0.1832	-3.0499	0.1489	0.5336
Random	7	0.7781	0.1273	-3.5139	0.0971	0.3779
Random	8	0.8813	0.0902	-3.9164	0.0583	0.2425
Random	9	0.9650	0.0643	-4.2714	0.0280	0.1230
Random	10	1.0000	0.0504	-4.5412	0.0139	0.0640
Uniform	2	0.2863	0.7851	-0.6389	0.5391	0.7474
Uniform	3	0.3782	0.5838	-1.2623	0.4255	0.9349
Uniform	4	0.4717	0.4210	-1.8124	0.3463	0.9601
Uniform	5	0.5693	0.2940	-2.2899	0.2886	0.9290
Uniform	6	0.6720	0.2279	-2.7285	0.2386	0.8550
Uniform	7	0.7781	0.1685	-3.1687	0.1858	0.7231
Uniform	8	0.8813	0.1194	-3.5509	0.1462	0.6080
Uniform	9	0.9650	0.0916	-3.8596	0.1217	0.5348
Uniform	10	1.0000	0.0765	-4.1101	0.1075	0.4951

Table E2. Index values from 27 simulated landscapes with equal patch proportions under three spatial arrangements and increasing number of patch types.

Spatial pattern	Number of patch types	Γ_1	Γ_2	RC_2
Aggregated	2	0.6435	-0.8954	0.3541
Aggregated	3	0.4782	-1.4642	0.3336
Aggregated	4	0.3953	-1.7808	0.3577
Aggregated	5	0.3246	-2.1228	0.3405
Aggregated	6	0.2802	-2.3472	0.3450
Aggregated	7	0.2371	-2.6176	0.3274
Aggregated	8	0.2187	-2.7266	0.3444
Aggregated	9	0.1949	-2.9390	0.3312
Aggregated	10	0.1722	-3.1039	0.3260
Random	2	0.4860	-1.3288	0.0415
Random	3	0.2865	-2.1546	0.0194
Random	4	0.1929	-2.7324	0.0145
Random	5	0.1398	-3.1806	0.0119
Random	6	0.1066	-3.5498	0.0094
Random	7	0.0847	-3.8560	0.0092
Random	8	0.0702	-4.1123	0.0112
Random	9	0.0622	-4.2863	0.0246
Random	10	0.0539	-4.4629	0.0309
Uniform	2	0.5188	-1.2474	0.1002
Uniform	3	0.3602	-1.9052	0.1329
Uniform	4	0.2792	-2.3193	0.1635
Uniform	5	0.2178	-2.7521	0.1450
Uniform	6	0.1515	-3.2563	0.0913
Uniform	7	0.1334	-3.4279	0.1192
Uniform	8	0.1172	-3.6777	0.1157
Uniform	9	0.0841	-3.9765	0.0951
Uniform	10	0.0769	-4.1027	0.1091

VITA

Bernard Ross Parresol was born in Washington, D.C., on September 15, 1953, the fifth child of Thomas and Rita Parresol. As a youth he was active in the scouting program and earned the Eagle Scout award. He received a varsity letter in cross-country track from Parkdale Sr. High where he graduated in June, 1972. A bachelor of science degree with honors in Forestry was conferred in June, 1977, from Michigan State University in East Lansing, Michigan. In October of 1977 he joined the forestry consulting firm of James M. Vardaman and Co. at their Alexandria, Louisiana, office. He served as Branch Manager of their Shreveport, Louisiana, office from February, 1979, to August, 1980. He entered Louisiana State University, Baton Rouge, Louisiana, in August of 1980 and completed a master of science degree in Forestry in May of 1983 and a master of applied statistics degree in December of 1983. Upon graduating he went to work as a Research Associate for the Department of Experimental Statistics, Louisiana State University. In July of 1986 he accepted a position as Mathematical Statistician with the Biometrics Unit of the USDA Forest Service, Southern Forest Experiment Station, in New Orleans, Louisiana. In 1988 the Biometrics Unit was given the designation of Institute for Quantitative Studies (IQS) because of its superior performance and productivity. In 1994 and again in 1997 he was an invited guest lecturer at Nanjing Forestry University in the city of Nanjing, People's Republic of China. Bernard Parresol married Lisa Barbé in May of 1995. In June of 1996 he accepted the position of Station Statistician in the new Southern Research Station, USDA Forest Service, in Asheville, North Carolina. He completed the degree of Doctor of Philosophy in Forestry at Louisiana State University in May, 1998.


DOCTORAL EXAMINATION AND DISSERTATION REPORT

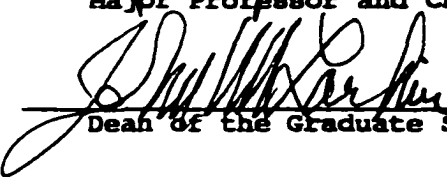
Candidate: Bernard Ross Parresol

Major Field: Forestry

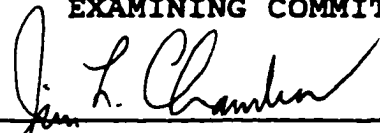
Title of Dissertation: Sample-based Forest Landscape
Diversity Indices

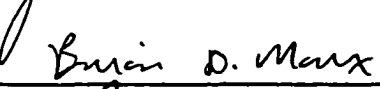
Approved:

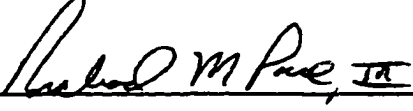

Major Professor and Chairman

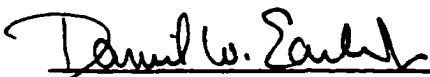

Dean of the Graduate School

EXAMINING COMMITTEE:


J. L. Chanton


Bruce D. Marx

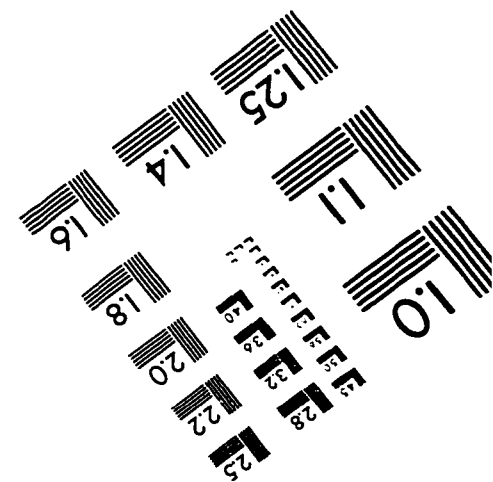
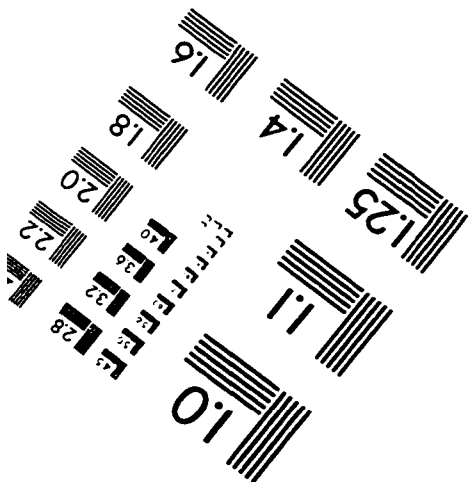
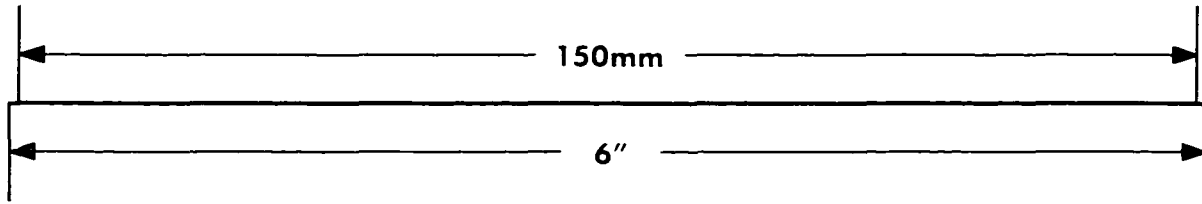
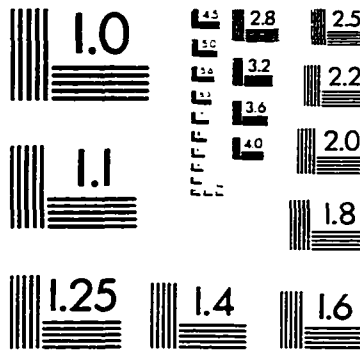
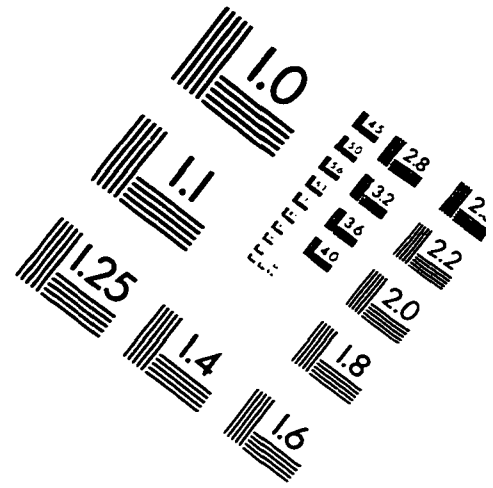
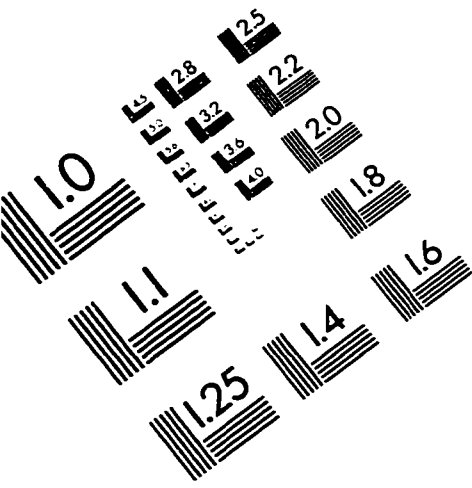

Richard M. Pace, III


Daniel W. Searles

Date of Examination:

April 9, 1998

IMAGE EVALUATION TEST TARGET (QA-3)



APPLIED IMAGE, Inc.
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved